

2014

University of North Carolina Wilmington
Master of Science in
Computer Science and Information Systems
Proceedings

<https://csbapp.uncw.edu/mscsis>

AN EVALUATION OF LONGITUDINAL FACE RECOGNITION PERFORMANCE THROUGHOUT
THE GROWTH AND DEVELOPMENT STAGES

Michael Kurt Sodomsky

A Thesis Submitted to the
University of North Carolina Wilmington in Partial Fulfillment
of the Requirements for the Degree of
Master of Science

Department of Computer Science
Department of Information Systems & Operations Management
University of North Carolina Wilmington

2014

Approved by

Advisory Committee

Li Hua

Jeffrey Cummings

Karl Ricanek, Jr.

Chair

Accepted by

Dean, Graduate School

TABLE OF CONTENTS

ABSTRACT	iv
ACKNOWLEDGMENTS	v
LIST OF TABLES	vi
LIST OF FIGURES	vii
CHAPTER 1: INTRODUCTION	1
Face Processing	1
Development and Aging Effects on Face	2
In The Wild Child Celebrity Dataset.....	3
Objectives and Contributions.....	4
CHAPTER 2: BACKGROUND	6
Face Recognition Techniques	6
Face Detection	7
Face Recognition Algorithms	7
Development and Aging Process.....	11
Evaluation Techniques	17
Aging Effects on Performance	18
Datasets.....	18
Evaluation Toolkits	21
CHAPTER 3: METHODOLOGY	23
CHAPTER 4: EXPERIMENTAL DESIGN	28
CHAPTER 5: RESULTS	37
CHAPTER 6: CONCLUSION	41
REFERENCES	43
APPENDICES.....	47

Appendix A: All-to-All Verification Results	47
Appendix B: Young-to-Old Identification Results	54
Appendix C: Augmented Young-to-Old Identification Results	61
Appendix D: Age Gap 4 - 3-5 to 9-11 Results	68
Appendix E: Experiment 4 - 3-5 to 12-14 Results	75
Appendix F: Experiment 4 - 6-8 to 9-11 Results	82
Appendix G: Experiment 4 - 6-8 to 12-14 Results	89
Appendix H: Experiment 4 - 6-8 to 15-17 Results	96
Appendix I: Experiment 4 - 9-11 to 12-14 Results	103
Appendix J: Experiment 4 - 9-11 to 15-17 Results	110
Appendix K: Experiment 4 - 9-11 to 18+ Results	117
Appendix L: Experiment 4 - 12-14 to 15-17 Results	124
Appendix M: Experiment 4 - 12-14 to 18+ Results	131
Appendix N: Experiment 4 - 15-17 to 18+ Results	138

ABSTRACT

An Evaluation of Longitudinal Face Recognition Performance throughout the Growth and Development Stages. Michael Kurt Sodomsky, 2014. Thesis Paper, University of North Carolina Wilmington.

Face recognition is the process of identifying or verifying the identity of an individual based upon the facial features of the individual. Automated face recognition has received a lot of attention over the last 15 years. The vast majority of research and development in automated face recognition has focused on adults. This work will evaluate the performance of automated face recognition of non-adults. Non-adults are those persons younger than 18 years of age. The changes of the cranium and thus, the outward appearance of the face, alters significantly from birth to approximately age 16. This period is composed of the growth and development stages. This work will investigate standard and commercial face recognition systems against a set of longitudinal images during the growth and development stages. It is an established fact that face recognition performance degrades against aging, i.e. when the enrolled face is temporally displaced by the probe face. This fact has been established in adult faces; however, the question has not been adequately investigated for non-adult periods, i.e. children from birth to adulthood. Our research has established this is a challenging problem

ACKNOWLEDGMENTS

I would like to acknowledge the following for their contributions and support throughout this work.

My friends Jeffrey Raynor and David Macurak for pushing me to complete this when it seemed overwhelming.

Ben Barbour for struggling with this work alongside me. You were there whenever I needed to bounce ideas off of.

Dr. Karl Ricanek for being so patient with me. He gave me inspiration to continue learning computer science past the undergraduate level. I would not be where I am today without him.

Finally I would like to thank my family for the constant support they provided throughout this work.

LIST OF TABLES

Table	Page
4.1 All-to-All Verification Experiment Dataset Usage	29
4.2 All-to-All Verification Matching Comparison	30
4.3 Young-to-Older Identification Experiment Dataset Usage	30
4.4 Young-to-Older Identification Matching Comparison	30
4.5 Augmented Young-to-Older Identification Experiment Dataset Usage	32
4.6 Augmented Young-to-Older Identification Matching Comparison	32
4.7 Growth Period Verification Experiment Dataset Usage	34
4.8 Growth Period Verification Matching Comparison.....	35
5.9 True Accept Rate at 1% False Accept Rate	39
5.10 True Accept Rate at 0.1% False Accept Rate	40
5.11 Cognite's FaceVacs True Accept Rate at 1% False Accept Rate	41

LIST OF FIGURES

Table	Page
1.1 Child Celebrities dataset examples.....	4
2.2 Example of Facial Features.....	8
2.3 Growth And Development Cranial Differences	12
2.4 Taylor, Karen T. Forensic art and illustration. CRC Press, 2010.[35]	16
2.5 ITWCC Ages of Images	20
2.6 ITWCC Age Ranges of Subjects in Years	20
3.7 In The Wild Child Celebrity Dataset.....	25
3.8 Normalized ITWCC images	27
4.9 Access Control Design	29
4.10 Photo Tagging Design	31
4.11 Augmented Photo Tagging Design.....	33
4.12 Growth Period Verification Design.....	36
5.13 Best and Worst ROC curves for All-to-All Verification.....	38
6.14 NIST Cognitec Results	42
6.15 Child Celebrity Dataset - ROC	48
6.16 Child Celebrity Dataset - DET	49
6.17 Child Celebrity Dataset - Score Histogram.....	50
6.18 Child Celebrity Dataset - CMC	51
6.19 Child Celebrity Dataset - ROC Scores	52
6.20 Child Celebrity Dataset - EER	53
6.21 Photo-Tagging Scenario - ROC	55
6.22 Photo-Tagging Scenario - DET.....	56
6.23 Photo-Tagging Scenario - Score Histogram	57

6.24 Photo-Tagging Scenario - CMC	58
6.25 Photo-Tagging Scenario - ROC Scores	59
6.26 Photo-Tagging Scenario - EER.....	60
6.27 Augmented Photo-Tagging Scenario - ROC.....	62
6.28 Augmented Photo-Tagging Scenario - DET	63
6.29 Augmented Photo-Tagging Scenario - Score Histogram.....	64
6.30 Augmented Photo-Tagging Scenario - CMC	65
6.31 Augmented Photo-Tagging Scenario - ROC Scores.....	66
6.32 Augmented Photo-Tagging Scenario - EER	67
6.33 Age Gap 3-5 to 9-11 - ROC	69
6.34 Age Gap 3-5 to 9-11 - DET	70
6.35 Age Gap 3-5 to 9-11 - Score Histogram	71
6.36 Age Gap 3-5 to 9-11 - CMC.....	72
6.37 Age Gap 3-5 to 9-11 - ROC Scores	73
6.38 Age Gap 3-5 to 9-11 - EER.....	74
6.39 Experiment 4 3-5 to 12-14 - ROC.....	76
6.40 Experiment 4 3-5 to 12-14 - DET	77
6.41 Experiment 4 3-5 to 12-14 - Score Histogram.....	78
6.42 Experiment 4 3-5 to 12-14 - CMC	79
6.43 Experiment 4 3-5 to 12-14 - ROC Scores.....	80
6.44 Experiment 4 3-5 to 12-14 - EER	81
6.45 Experiment 4 6-8 to 9-11 - ROC	83
6.46 Experiment 4 6-8 to 9-11 - DET	84
6.47 Experiment 4 6-8 to 9-11 - Score Histogram	85
6.48 Experiment 4 6-8 to 9-11 - CMC.....	86
6.49 Experiment 4 6-8 to 9-11 - ROC Score	87

6.50	Experiment 4 6-8 to 9-11 - EER	88
6.51	Experiment 4 6-8 to 12-14 - ROC.....	90
6.52	Experiment 4 6-8 to 12-14 - DET.....	91
6.53	Experiment 4 6-8 to 12-14 - Score Histogram.....	92
6.54	Experiment 4 6-8 to 12-14 - CMC	93
6.55	Experiment 4 6-8 to 12-14 - ROC Scores.....	94
6.56	Experiment 4 6-8 to 12-14 - EER	95
6.57	Experiment 4 6-8 to 15-17 - ROC.....	97
6.58	Experiment 4 6-8 to 15-17 - DET.....	98
6.59	Experiment 4 6-8 to 15-17 - Score Histogram.....	99
6.60	Experiment 4 6-8 to 15-17 - CMC	100
6.61	Experiment 4 6-8 to 15-17 - ROC Scores.....	101
6.62	Experiment 4 6-8 to 15-17 - EER	102
6.63	Experiment 4 9-11 to 12-14 - ROC	104
6.64	Experiment 4 9-11 to 12-14 - DET.....	105
6.65	Experiment 4 9-11 to 12-14 - Score Histogram	106
6.66	Experiment 4 9-11 to 12-14 - CMC.....	107
6.67	Experiment 4 9-11 to 12-14 - ROC Scores	108
6.68	Experiment 4 9-11 to 12-14 - EER.....	109
6.69	Experiment 4 9-11 to 15-17 - ROC	111
6.70	Experiment 4 9-11 to 15-17 - DET.....	112
6.71	Experiment 4 9-11 to 15-17 - Score Histogram	113
6.72	Experiment 4 9-11 to 15-17 - CMC.....	114
6.73	Experiment 4 9-11 to 15-17 - ROC Scores	115
6.74	Experiment 4 9-11 to 15-17 - EER.....	116
6.75	Experiment 4 9-11 to 18+ - ROC	118

6.76	Experiment 4 9-11 to 18+ - DET	119
6.77	Experiment 4 9-11 to 18+ - Score Histogram	120
6.78	Experiment 4 9-11 to 18+ - CMC	121
6.79	Experiment 4 9-11 to 18+ - ROC Scores	122
6.80	Experiment 4 9-11 to 18+ - EER	123
6.81	Experiment 4 12-14 to 15-17 - ROC	125
6.82	Experiment 4 12-14 to 15-17 - DET	126
6.83	Experiment 4 12-14 to 15-17 - Score Histogram	127
6.84	Experiment 4 12-14 to 15-17 - CMC	128
6.85	Experiment 4 12-14 to 15-17 - ROC Scores	129
6.86	Experiment 4 12-14 to 15-17 - EER	130
6.87	Experiment 4 12-14 to 18+ - ROC	132
6.88	Experiment 4 12-14 to 18+ - DET	133
6.89	Experiment 4 12-14 to 18+ - Score Histogram	134
6.90	Experiment 4 12-14 to 18+ - CMC	135
6.91	Experiment 4 12-14 to 18+ - ROC Scores	136
6.92	Experiment 4 12-14 to 18+ - EER	137
6.93	Experiment 4 15-17 to 18+ - ROC	139
6.94	Experiment 4 15-17 to 18+ - DET	140
6.95	Experiment 4 15-17 to 18+ - Score Histogram	141
6.96	Experiment 4 15-17 to 18+ - CMC	142
6.97	Experiment 4 15-17 to 18+ - ROC Scores	143
6.98	Experiment 4 15-17 to 18+ - EER	144

CHAPTER 1: INTRODUCTION

Facial recognition has many characteristics that make it a desirable biometric modality. These characteristics include universality, acceptability, semi-permanence, and easy collectability[17]. Universality is a character trait which most or all people should possess, such as an iris or a fingerprint. Acceptability is an individual's comfort with collection methods being presented, i.e sticking the subject's eye up to a sensor is less acceptable than using a fingerprint scanner. Semi-permanence is very important for longitudinal work. This characteristic is the rate of change of the feature being collected with slow changing features being better for face recognition. Easy-collectability is the characteristic in which the sensor collecting the feature is non-intrusive to the individual. The face is a very strong candidate for all of the features above, which makes this a highly researched topic.

The proposed experiments will leverage the Child Celebrities dataset. This dataset, which was developed for this work from publicly sourced images on the Internet, contains longitudinal face data of children and young adults. An example of the images in this dataset are shown in figure 1.1. Face recognition is a challenging problem, which has not been thoroughly investigated across human aging. Aging in respect to facial recognition, is the changes within the cranium and face that manifest changes in the image. These temporal changes will cause performance degradation over a period of time. Mathew Turk stated that "developing a computational model of face recognition is quite difficult, because faces are complex, multidimensional, and meaningful visual stimuli[38];" however, when aging information is added to this problem, it becomes infinitely more challenging. Evaluation of this problem will be completed by running multiple baseline and commercial algorithms against a novel dataset developed for this work. This thesis examines the current state of facial recognition algorithms with a focus on the growth and development stage of humans.

1.1 Face Processing

Throughout this thesis the term 'face processing' will be used to generalize all techniques used to process and understand a face image. This includes but is not limited to face recognition, face-based soft biometrics, age-estimation, sex-determination, etc. This work will focus on face recognition; although,

the novel dataset and insights rendered from this work can illuminate other areas of face processing.

Face recognition is defined as the following: given a still or a video, the system must identify or verify one or more persons' within the given scene against a stored database of face template(s). Face recognition has two modes of operations, identification and verification. Identification is a one to many match where an individual is matched against an entire gallery of individuals. Verification is the process in which a subject is matched against one subject. Identification answers the question "Who am I?", where verification answers "Am I this person?" Identification can be seen as verification applied over the set of all enrolled faces.

The identity of a person is not the only usable information that can be extracted from the face. Soft biometrics are "characteristics that provide some information about an individual, but lack the distinctiveness and permanence to sufficiently differentiate any two individuals[18]." All of the information gathered from this face processing technique is highly valuable for researchers, security professionals and even marketers as identity may not be relevant information. This is because the person being analyzed may not exist in the known set. For example, in the Boston Marathon Bombing, April 15, 2013, the investigation wanted to find the suspects that may have planted the improvised explosive devices. Providing identity using traditional face recognition does not have the investigational power at the early stages of investigation. Once a list of possible perpetrators were deduced from the video/still image data, identity could have been used to locate contact information from DMV data, etc. Soft biometrics can provide "rich information about a person: the size and geometry of the chin, lips, nose, eyebrows, and other face components can be used to distinguish gender, race, and ethnicity, while creases, lines, sagging and wrinkles can reveal clues about age[19]."

1.2 Development and Aging Effects on Face

The face is a dynamic canvas which changes often throughout a person's life. As a person develops and ages many processes occur which can alter the appearance of the face. Development and aging is a cumulative effect from both facial textural wrinkling of the skin and changes in the underlying 3-D structure of the cranium[11]. In the context of facial processing, aging is the biological changes that

manifest in changes in digital images over a period of time.

Changes in the cranium shape and skin texture(wrinkling and sagging) are different for each person as they age; however, for majority of the populations these aging trends often occur at a particular year. "Cranial shape changes are the predominant changes in children and young adults, especial from birth to age 16, while in later years, reduced muscle tone, collagen, and skin wrinkling and sagging alters the appearance of the person[31]." Photo-aging is the process in which sunlight alters the face and expedites aging. People with fair skin or high amounts of sun exposure are particularly at risk for this type of aging. Some other external factors also alter the appearance of the face. These factors include "ancestry, gender, health and disease, tobacco and drug use, diet, stress-related sleep deprivation, bio-mechanical factors, gravity, and hyper-dynamic facial expressions[31]."

1.3 In The Wild Child Celebrity Dataset

The In the Wild Child Celebrity, or ITWCC, dataset was created to assist in the evaluation of facial recognition tools over a longitudinal span, specifically over the non-adult growth and development stages. The In the Wild Child Celebrity dataset contains over 1700 images, with the gender ratio being approximately 50-50. Each subject has at least three images within this corpus and at least two of these images are of differing ages which are younger than the age of 16. An example of a subject in the In The Wild Child Celebrity dataset is shown in figure 1.1. The dataset was collected by gathering images on publicly available websites of individuals, whom have high media exposure. ITWCC contains images of child celebrities over their career.

Many data corpses exist in the space of facial recognition; however, very few contain images spanning multiple years of each subject. FG-NET is one of the few datasets in which longitudinal data contains people younger than 16, but very few subjects, only 82 usable, are present in this dataset[24]. An additional dataset used in this paper is the LFW data corpus[16]. LFW, Labeled Faces in the Wild, is a data set of unconstrained face images. unconstrained face images contain a "large amount of variation in pose, lighting, expression, background, race, ethnicity, age, gender, clothing, hairstyles, camera quality, color saturation, focus and other parameters[16]." LFW contains over 13000 images, with 5749 subjects



(a) age 4



(b) age 16

Figure 1.1: Child Celebrities dataset examples

and 1680 subjects that have multiple images. The Chinese Twin Dataset is also used in this work[32].

1.4 Objectives and Contributions

In this work, a challenging dataset is introduced and face recognition performance was evaluated on non-adult faces. This work will provide an evaluation methodology which can be performed to evaluate longitudinal data and an evaluation of commercial and public domain algorithms on a multitude of datasets. The following metrics are used: verification rate against false matching aka Receiver Operating Characteristic(ROC) curve and Rank-accuracy. Due to the small number of subjects in the In The Wild Child Celebrity dataset the identity experiments will be augmented with the LFW dataset and Chinese Twins dataset. These datasets were selected because they mirror the acquisition methods used for the In The Wild Child Celebrity dataset, i.e. uncontrolled data capture. This work aims at uncovering if any problems exist with the facial recognition space against longitudinal data. With the increase of facial recognition in todays society aging will become a larger and larger problems for algorithms to handle.

This thesis is organized as follows: Chapter 2 is an exploration of previous work in the work of face recognition and child aging. Chapter 3 is an explanation of the design of each of the experiments performed in this work. Chapter 4 will be the results of each experiment and will discuss the relevance of each. Chapter 5 is a summary of the work in this paper and finally chapter 6 will discuss future work. This thesis examines the current state of facial recognition algorithms with a focus on the growth and development stage of humans.

CHAPTER 2: BACKGROUND

Facial recognition is a challenging topic that has been researched very heavily recently; however, human aging, in respect to algorithm performance, has not been evaluated thoroughly enough. The largest issue with the subject is the vast amount of data that is needed to process to fully understand the face. Growth, development and aging factors affect how well our methods perform on the face over time. As the face changes over time, the ability to recognize the person becomes challenging. Human perception is often studied as clues to how we can fundamentally understand one's perception of another age. Humans are very accurate at judging the ages of people, specifically between the ages of 5-70[2]. Mark et al. evaluated human correlation between general craniofacial remodeling of the face and development of wrinkles[27]. In his experiments, they created a scenario in which the shape of the face would change and wrinkles would be applied to the different shapes of the face. They concluded that craniofacial development and wrinkling does not necessarily correlate to age information, but that the relationship between the two contain the age information.

2.1 Face Recognition Techniques

Recognizing the face within an image is a multi-step process. Object detection is the first part of this process and is often considered the most difficult step. Feature extraction occurs once a face has been found within the image and results in usable data. Facial features vary on the classification method used which can range from the components of the face such as eyes, nose and mouth to abstract features such as lines and fiducial points[40]. Finally, the matching process uses the data from the feature extraction stage to verify or identify the face against a database of face templates. A fully automated system will accomplish all of these steps. If any of these steps fail the face is not recognized.

Object detection in terms of computer vision is *finding* an object or class in a digital image. Face detection is a branch of object detection in which the classifier is trained to detect only faces. Within an image, any number of faces must be found and segmented from the background. Due to the sheer amount of data in an image, the detection algorithm must be robust enough to discover the object. Pose, illumination, distance from the camera, and noisy backgrounds make this a difficult problem. Object

detection provides the location and size of an object within an image, commonly known as a bounding box. Furthermore, it does not provide any detailed information about the object. For example, if the detector was designed to locate an apple in an image, it would be able to discover the apple, but it would not give you information about its ripeness or color. Two common implementations of the face detection process are below.

2.2 Face Detection

Paul Viola and Michael Jones created an algorithm, commonly known as Viola-Jones, which uses cascading Haar-like classifiers to detect faces[39]. The system was designed to detect faces very rapidly by using an "integral image." The integral image is a summation of pixels in a region that is used for classification which bears resemblance to Haar-like objects. This algorithm was trained using a modified Adaboost which focuses on grouping critical features. The dominant features of the face would become a weak classifiers for detection. A weak classifier is a feature that is needed to identify a part of an object, but it does not have enough discriminatory information for it to identify the entirety of the object. An example of a weak classifier might be the handle on a coffee mug. The weak classifier will show that the handle exists, but that doesn't necessarily mean the object in the image is a coffee mug. Viola-Jones uses a cascade of these weak identifiers in series to eventually prove that a face is found. In figure 2.2, three possible weak classifiers are shown. In this example if the eyes are found in the red weak classifier, the algorithm moves onto the green and finally the blue. If any of these regions is not found within the object, detection fails; however, if each subsequent weak classifier passes its evaluation, the object will be detected.

This work will leverage the usage of a fully automated face recognition system, Cognitec's FaceVacs[10]. Additionally OpenCV's(Open Source Computer Vision)[9] face detector will be used in all non-commercial algorithms in the following experiments. OpenCV's face detector is based on a version of Viola and Jones's algorithm(Viola-jones)[39]. While openCV's source code is open-sourced for the community, Cognitec's algorithms are private and not known.

2.3 Face Recognition Algorithms

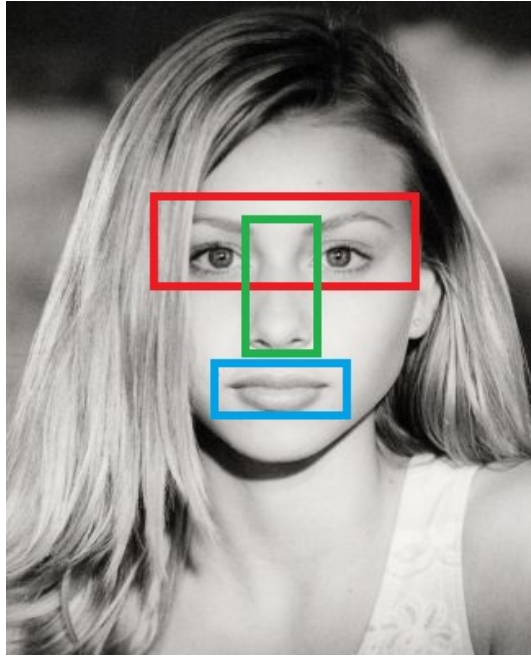


Figure 2.2: Example of Facial Features

A face recognition algorithm is the full process of accepting an image in and matching the image to a template to determine identity. This is completed by registering an image, extracting the information from the input, and then finally matching the input to a stored face template which will produce a result which can be evaluated. The feature extraction process of face recognition is very important for accurate representations of the data. Feature extraction methods can vary from methods based on lines, edges and curves, facial feature based methods, or geometrical methods[40]. The face contains shape, textural and other information which must be captured in this step to form a representation of the image. In this step of the process, factors other than the face itself can influence the data. Lighting and occlusions, such as glasses or hats, can mask parts of the face. Any Aging between the stored template and the new input image will now become entered into the system. Methodologies for feature extraction varies greatly depending on the final goal. Holistic matching methods, feature-based matching methods, and hybrid methods are used to accomplish this task. Holistic matching methods take the entire face image as input for the system. Feature-based matching methods take a local feature from the image such as the periocular region, mouth or nose. A hybrid method would be a combination these two extractor types to

further mimic the human recognition process. Once these features have been acquired the features then need to be compared using some sort of metric.

Principle Component Analysis(PCA) is a common method in facial recognition to extract useful information from an image. PCA is a statistical method that evaluates a set of data and uses an orthogonal transformation to produce a set of linear variables known as principle components(eigenvalues and eigenvectors). PCA simply finds the most *important* information within an image and in most cases for facial recognition, these components will describe a face. Turk and Pentland developed an algorithm for facial recognition that leverages PCA known as Eigenfaces [37]. Turk's goal with this system was to provide a fast, simple and accurate method for recognition. The system is set up by first acquiring a set of face images for training. PCA is used to transform the face images into a set of principle components, which then can be repeated for new images[37]. The eigenfaces, which is the set of eigenvectors, of the training information produce a subspace known as the *face space*. Once the training of algorithm is completed, whenever a new image is introduced to the system it can be projected into the *face space* to see whether it is a face and if it is recognizable. The comparison method used in the algorithm is a simple euclidean distance measurement within the projected *face space*. Eigenface algorithm is often considered very sensitive to unconstrained image sets.

Phillips et al. released a baseline algorithm for evaluation of the Good, the Bad and the Ugly Face Recognition Challenge(GBU)[29]. Local Region Principle Component Analysis(LRPCA) was developed for this task and provides refinements to the standard PCA algorithm. According to Phillips et al., considerable performance refinements which extended PCA were, "representing a face by local regions, a self quotient normalization step and weighting eigenfeatures based on Fischer's Criterion"[29]. LRPCA was designed for the GBU dataset, thus the resulting algorithm is designed to handle structured frontal faces. The algorithm begins by extracting the face from the image and geometrically normalizing the face. It is scaled, cropped and rotated so the eyes are in the same position in each image. Once the face normalization has been completed the algorithm crops out thirteen specific regions which will be used for recognition. These regions have been preselected by the authors and contain large amounts of

varying data, such as eyes, nose and mouth. After the 14 images(full face and thirteen face regions) have been cropped the images are finally normalized for PCA. The images are projected into 14 different subspaces so each region can be matched separately. A template is generated in this system by calculating 250 coefficients from each of the regions to produce a single template for each face. The PCA subspaces are scaled and then weighted based upon a Fischer Criterion to push different people farther apart in the space. When a new face is introduced to the system, it uses the same steps as above to produce a template. The new template is compared to a already computed template by measuring the similarity through the computation of the Pearson's correlation coefficients.

Linear Discriminant Analysis(LDA) is a statistical method to separate multi-class data which can be used for classification. The Fisherface algorithm, created by Peter Belhumeur et al., uses Linear Discriminant Analysis(LDA) to accomplish facial recognition[5]. The algorithm was proposed as a way to handle illumination and pose variance. Fisherface is also a linear projection based algorithm, similar to Eigenface; however, while Eigenface maximizes the separation of all classes within the projected space, LDA maximizes the between-class separation and minimizes the within-class separation. LDA's min-max separation reduces the effects of illumination, expression, and eye wear by clustering within-class individuals. In the experimental results produced by Belhumeur et al., it was shown to have lower error rates; however, the dataset used in this experiment was very small[5].

The Spectrally Sampled Structural Subspaces Features(S4F) algorithm[20] is an openly sourced statistical learning-algorithm used in the OpenBR evaluation system[23]. The algorithm was originally designed to study demographics[22] and aging[21] within the facial recognition space. The algorithm begins by using the OpenCV Viola-Jones object detector[39] for face and then a c++ port of the ASEF eye detector[7] to begin the registration process. The registration process uses a rotation and affine transformation for scaling and then completes an illumination normalization step. This algorithm uses both LBPs[3] and SIFT[25] as descriptors. Histograms of the LBP and samples of the SIFT are taken from local regions within the face to be input for a PCA decomposition. The next step uses weighted spectral sampling in which all feature vectors from the previous step are concatenated and randomly

sampled. This sampling is performed based upon the variances of each dimension and then LDA is calculated for each region. The samples are reconcatenated again and are normalized. The matching of templates in the S4F algorithm is completed by using a L_1^{byte} distance metric[23]. The S4F algorithm was the first openly sources algorithm that was submitted for NIST FRVT tests[30]. The algorithm "ranked 13th with a true accept rate of 64.8% on mugshots and 14th with a true accept rate of 76.1% on visas each at a false accept rate of 0.1%[23].

The CohortLDA, formerly known as LDAIR, baseline algorithm which evaluates the color space for facial recognition[26]. The algorithm leverages the RGB color space as well as the YIQ color space. Both the red channel and I channel are used for calculations; however, with illumination changes the red does not work effectively on its own. The problem was resolved by taking the I channel into account as the combination of both color channels will be less invariant to smaller lighting gradients. There still remains a problem with large lighting changes within the image, but this was fixed by using a logarithm transformation and then applying a zero-mean and standard deviation normalization. Once images have been normalized, a Fisher LDA is applied to the data. The LDA calculations create a large dimensionality problem which needs to be reduced. CohortLDA uses a PCA to further reduce the dimensions. After the system is trained the Cohort normalization process is used to define match thresholds. The cohort normalization process attempts to adjust the match verification threshold for each *cohort set* of images. The CohortLDA algorithm was tested against other algorithms including their LRPCA algorithm[29] and the this algorithm was proven to be a robust algorithm. When compared to other non-commercial algorithms this proved to be superior except for the FRVT 2006 fusion algorithm.

2.4 Development and Aging Process

The face develops and ages in numerous ways which pose challenges for face processing techniques. Humans have the ability to recognize a person from years ago; however, the person does look fundamentally different. These differences will affect facial recognitions algorithm's ability to match individuals over periods of time. Natural growth, environmental perturbations and physiological stresses affect the development and aging of the face[4]. Craniofacial, the combination of the cranium and face,



Figure 2.3: Growth And Development Cranial Differences

development and aging rates vary across a persons lifespan. Craniofacial changes generally manifest in shape changes in younger years and texture deformations in elderly[27].

Childhood craniofacial development is almost solely comprised of cranial growth. As you can see from figure 2.3, The skull on the left is a child while the skull on the right is a matured adult. These underlying skeletal changes will alter the appearance of the face. Cranial growth will not greatly change the features within the face, but they will change the proportions between the major features. For example, a growth period of the jaw will cause the face to elongate, but will not alter the individual features shapes. Throughout developmental changes the features of the face will remain alike to their original. This growth pattern is known as gnomatic growth[35]. The eyes, nose and mouth will alter, but the growth of the underlying structure will change the geometrical distances between them more then the feature itself. These distance changes from the underlying skeletal growth will alter the appearance of the face significantly.

It's important to note that the highest rates of craniofacial change is in the childhood years. We this age range to be the most challenging for facial recognition systems matching performance. Factors such as puberty and the growth of permanent teeth affect craniofacial growth. Maturation is achieved in males between the ages of 12 and 15 while between 10 and 13 in females[12]. At this age the underlying structure of the face will not grow as rapidly as before, but changes will continue to occur.

The underlying structure of the face is the largest discriminate factor in the growth and development stages of children. As expected in childhood development, there are large jumps in growth at certain periods, i.e. growth is non-linear. The most significant growth periods in children are between birth and 1 years old and from age 1 to 5[12]. Farkas et al. developed an experiment in which growth patterns

of the skull were tracked in children over a period of time. Seven measurements were taken from the face across 1594 Caucasian Americans in this study[12]. The seven measurements in this study are: Face Height, Upper Face Height, Mandible Height, Face Width, Mandible Width, Middle Third Face Depth, and Lower Third Face Depth. These growth areas measured were used to show the growth differences in particular regions of the cranium.

In Karen T. Taylor's book "Forensic Art and Illustration" she describes the changes of the cranium and face year by year from childhood to young adulthood[35]. This book is written for forensic artists, but the information is just as valuable for computer science and other disciplines.

Age 1½

- the head is very rounded and the facial contours are full and rounded
- the face occupies the lower portion of the head and there is a lot of cranium
- the bridge of the nose is flattened
- there is the appearance of epicanthic folds and the eyes look large and rounded
- the hair is soft and fine



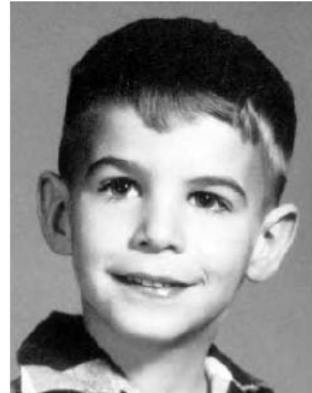
Age 3

- the cranium expands to accommodate the growing brain
- the eyes are slightly more elongated and less rounded
- the maxilla and mandible have enlarged and widened to allow room for the deciduous dentition
- the interorbital distance is almost established



Age 4

- the nose is still small and buttonlike with small nostrils, but the bridge has begun to form
- the interorbital distance is basically established
- the anterior "baby teeth" are visible
- the ears seem very large and very low on the head
- the chin has taken some shape



Age 5

- the bridge of the nose continues to rise up, lifting some of the excess skin from the medial corners of the eyes
- the face continues to elongate as the nose length and the chin length increase
- growth pattern of the hair seems firmly established, it is less fine, and the color darkens



Age 6½

- the forehead has become less prominent and bulbous looking
- the bridge of the nose continues to rise up and the nostril size and nose width increase slightly
- the nose continues to grow in length, as does the chin
- the forms of the lower cartilages of the nose become apparent and the tip takes shape
- the central and lateral deciduous incisors have been lost and replaced by the permanent maxillary central incisors, thus the dentition is “mixed”
- the mouth has to grow to accommodate the permanent teeth

*Age 8*

- there is not much proportional change
- the permanent maxillary lateral incisors have appeared but they are not fully down

*Age 9½*

- the bridge of the nose continues to rise up and the nostril size and nose width still increase slightly
- the face elongates slightly
- the squarish form of the chin becomes obvious
- the permanent maxillary lateral incisors are fully down
- the teeth seem big for the face

*Age 11*

- the “childlike” face is looking more “juvenile” as some of the facial forms become more apparent due to less “baby fat”
- the bridge of the nose continues to rise up
- the upper lip has remained about the same for several years
- the form of the chin becomes more masculine and square
- the ears still seem large for the face although they do not appear so low on the head



Age 12

- the nose continues to grow, both in the bridge and the nostril size
- the teeth still seem big for the face
- the permanent maxillary canines or cuspids are in
- the mandible and chin continue to grow
- the neck musculature is slightly more masculine



Age 15

- the “juvenile” face is starting to look more “teen-aged” (and would probably be even more noticeable in a female)
- the teeth don’t seem so oversized for the face as the mandible continues to grow and become more masculine
- the cheekbones have become relatively more prominent
- the ears don’t seem so oversized for the face as it has elongated and they don’t seem too low
- the eyebrows have become more masculine
- oops! ... a zit or two!



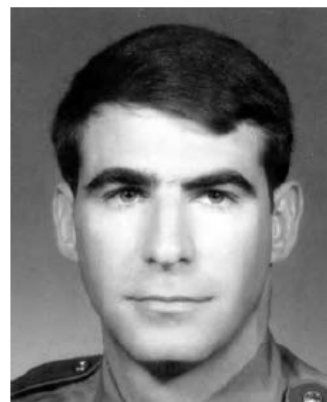
Age 20

- the nose has grown even more, revealing the nasal bones at the bridge
- the forehead shape has remained consistent though it has grown and risen up at glabella
- the face now looks mature, particularly due to the appearance of facial hair
- the eyebrows are even fuller
- the angle of the mandible is much more squared and masculine
- the neck musculature is mature



Age 25

- the eyebrows are heavier
- the mandible has continued to square and form, looking more masculine and mature
- the neck musculature is more defined and the Adam’s apple is visible



In early adulthood, ages 16-30, minor changes continue to effect the face. The rate of changes within this period of life are not as rapid as the childhood years. The focus of change in this period of development begins to more prominently effect the soft tissues of the face, but skeletal changes continue to occur. Early adulthood shows the first signs of soft tissue stressing. Hyper-dynamic expressions will start to show wrinkles on the face. Fine facial lines will appear horizontally on the forehead, vertical lines between the forehead and thin lines around the outside corners of the eyes will appear[4]. The cranium continues to slightly grow in this period of development, such as mandible length, nose bridge, as well as other underlying structures[4, 35]. Around 25 years of age the cranium ends its phase of ossification, or bone growth[35]. With the end of ossification, the cranium will start to change form with age instead of growing.

While majority of changes of a youthful face manifest with cranial shape changes solely, the adult face's soft tissues are the predominant changes. From ages 40 to 50 there are noticeable changes to skin texture while minimal changes are found in younger years[33]. The aging rate of adults differs heavily on the individual, which is not the case for non-adults[4]. These differences can be attributed to genetics and external features. Biological changes in adults alters the shape and texture of a face. As the skull continues to change with age, the eyes appear smaller as they sink in deeper into their orbits. Noses and ears can lengthen as cartilage continues to grow through life. As the skins elasticity begins to degrade wrinkles form, more notably in the eyelids, and the corners of the mouth[?]. These feature begin to sag and change in size thus changing the relationships of the features of the face.

2.5 Evaluation Techniques

All biometric systems can be evaluated in the same way. Fundamentally each system will match at least two biometric templates, one being the stored template and the other being the new users, to produce a score which will determine acceptance or rejection. This match score is a standardized number that shows the likeness between two templates. Both genuine users and impostors are used to evaluate a system. Ideally all genuine users should be accepted while all impostors should be rejected. Important metrics to note are as follows: true accept rate(TAR), true reject rate(TRR), false accept rate(FAR), and

finally false reject rate(FRR). True accept rate is the ratio of genuine users whom have been accepted, while the true reject rate is the the ratio of impostors who have been correctly rejected. The false accept rate is the ratio of genuine of users whom are rejected and finally the false reject rate is the ratio of impostors who gain access to the system. A user is rejected or accepted by comparing the match score to a match threshold. The match threshold is an arbitrary number that each system is tuned to achieve the results it requires.

2.6 Aging Effects on Performance

Brendan Klare and Anik Jain evaluated multiple facial recognition algorithms on a time lapsed database[21]. In this study they introduced an algorithm, Random Subspace Linear Discriminant Analysis(RS-LDA). The data used in this experiment is mugshot data collected from Pinellas County, Florida. The set of images used for the experiment was over 200,000 and was divided into multiple segments. The divisions of dataset was divided by measuring the time between the initial photograph taken and any repeat images of that subject. The divisions are as follows: zero to one year time difference, one to five year time difference, five to ten year time difference and more than 10 years difference. 8000 images were used for the training sets for each of the divisions, except for the 10 or more, to create a total of 5 systems which would be evaluated. The evaluation used both Cognitec's FaceVacs[10] and Pittpatt[15] for the commercial based benchmark algorithms in this experiment. Their initial experiments used the commercial systems to show a degradation in system performance as the age divisions became more difficult. The commercial algorithm's results were masked in order to keep anonymity of their performance. The best performing algorithm had a 96.3%, 94.3%, 88.6%, and 80.5% true accept rate at 1.0% false accept rate in the following time lapse periods respectively: 0 to 1, 1 to 5, 5 to 10, 10+. The experimentation concluded that the performance does decrease as the time between probe and gallery increases. Furthermore, they also showed that by training a system for non-aging scenarios on aging data will lower performance and that training a set of data for a particular age range will yield the best performance.

2.7 Datasets

The *In The Wild Child Celebrity dataset* is the largest longitudinal dataset that has been devel-

oped specifically for the non-adult growth and development stages. This data corpus focuses on having large sets of individuals where the subjects development can be observed. The faces in this dataset are unconstrained, which means that the images are not captures with the same camera, lighting or pose of the subject. This data corpus is designed to emulate a real-life capture scenario. Images were captured by exploiting the fame of the subjects and gathered through public media sources. The data was captured until December of 2013. The requirements used to develop this dataset are as follows:

1. The subject must have at least three images to qualify.
2. The subject must have at least two images less than 16 years of age.
3. The date that the photo was taken must be available.
4. Only images with frontal facing images are used.
5. Images must only contain one face.

In addition to the age information, other meta-data is also captured. Age, race, gender, data of the photo, name, a unique photo identifier, and a conditional makeup and glasses marker, and the URL of the image is recorded for each entry. This information can further illuminate the differences in gender specific aging variation's and occlusion's effects on facial recognition systems.

The In The Wild Child Celebrity dataset is compromised of 304 subjects and 1718 images. The subjects ages within this dataset range from 5 months to 32 years. The dataset contains 872 female subjects and 846 male images. The average age of all images is 13.418 years with a standard deviation of 3.432 years. The average age of the subjects for acquisition into this dataset is 10.2119 years with a standard deviation of 3.986 years; furthermore, the average age of final capture is 16.265 years with a standard deviation of 4.467 years. Figure 2.5 shows the number of images for each age. Figure 2.6 expresses the number of subjects with a particular longitudinal age range, i.e. the maximum age less the minimum age. Due to the nature of the images being "in the wild", 516 of the images are marked with the makeup tag and 43 with the glasses tag.

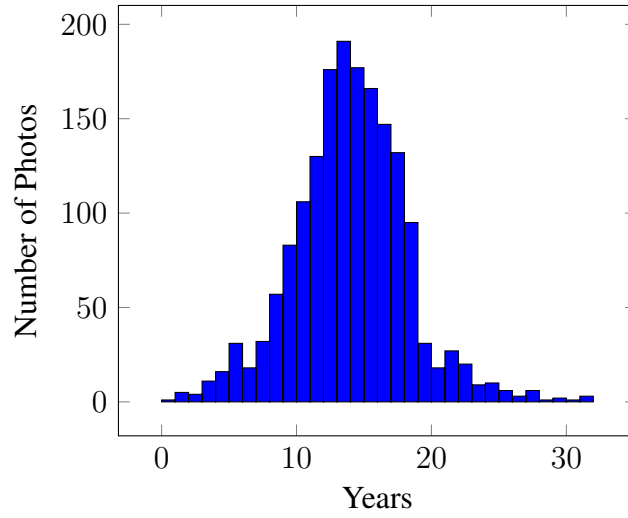


Figure 2.5: ITWCC Ages of Images

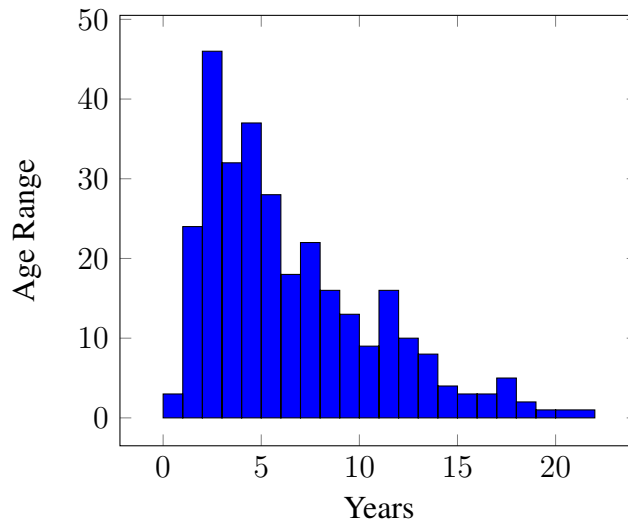


Figure 2.6: ITWCC Age Ranges of Subjects in Years

The LFW, Labeled Faces in the Wild, is used to in particular experiments to augment the In The Wild Child Celebrity dataset[16]. Similarly to the In The Wild Child Celebrity dataset, LFW contains completely unconstrained images. The background, lighting, camera all have variances throughout the images. This dataset was designed to be used with the unconstrained matching problem. The dataset is split into a training set and testing set to target this problem further. The people in each set are mutually exclusive to each other, which allows for no model to be built for a certain subject at training time[16]. The The LFW contains 13,233 images with 5749 unique subjects. All of the images in this dataset are 250x250 pixels and mostly in color. 1680 people in this corpus have multiple images. The Viola-Jones face detector ran through every image in this dataset to make sure a face was present in each image. Each face is centered and cropped within the 250x250 pixel image.

Additionally, the CASIA Multi-modal Database of twins will be used to augment the In The Child Celebrity dataset[32]. The database of images collected was developed to explore the impact on identical twins on biometric systems. This dataset is a combination of face, iris, and fingerprint modalities. The face database contains 134 subjects and there are approximately 20 images for each subject. The images were all taken from a USB camera in color; however, like the other databases the background and illumination of the images are unconstrained. The images are all size 480x640 pixels; however, the face can range from 280x300 to 300x400 in each image. The images within the face database were taken in rapid succession so longitudinal information is not present.

2.8 Evaluation Toolkits

Biometric Evaluation Systems are a good way for researchers to uncover information about individual algorithms. Colorado State University has released an Evaluation system for face recognition systems to compare against. CSU released a set of baseline algorithms and a methodology for comparison. The baseline algorithms that where released are as follows: a standard PCA, a combination PCA and LDA algorithm[13], a Bayesian Intra-personal/Extra-personal Image Difference classifier[36] and an Elastic Bunch Graph Matching Algorithm[8]. In 2011, two new baseline algorithms where released which were shown to work significantly better. This new baseline system included LRPCA[29]

and LDA with color spaces and cohort normalization[26]. The new baseline's were designed to be used with The Good, The Bad, and The Ugly Face Recognition Challenge to be a "strong lower bound to performance"[29].

Another biometric evaluation toolkit and API is Open Source Biometric Recognition(OpenBR) collabroratory[23]. OpenBR is a collaborative tool that provides a method for researchers to compare algorithms in a controlled environment. Klontz et all stated that "OpenBR provides tools to design and evaluate new biometric algorithms and an interface to incorporate biometric technology into end-user applications"[23]. The system introduces a new language for image processing and algorithm development. This *algorithm description* defines the template enrollment and comparison methods for each algorithm. Instructions to control detection, normalization, representation, extraction and matching can all be changed within this algorithm description. The system comes with the Spectrally Sampled Structural Subspaces Features(4SF) algorithm when installed[20]. OpenBR leverages OpenCV for its already established algorithms, such as Eigenface[37], Fisherface[5], and LBP[3], and extensive support from the open source community. The system contains a simple command line interface which can be used to create and evaluate the algorithms, as well as a API for creating plug-ins. The system's recognition evaluation tool provides a set of charts which can be seen within appendix A - N.

CHAPTER 3: METHODOLOGY

This thesis will evaluate the challenges of using facial recognition techniques on adolescent faces. The non-adult face goes through many structural and textural changes throughout the growth and development stages which pose problems for many systems which require face input. Passport identification, photo-tagging, law enforcement, and other biometric related tasks use facial recognition to determine identity; however, due to the large ranges of changes this is very challenging. The objective of this thesis will be to create scenarios which the In The Wild Child Celebrity(ITWCC) and additional datasets can be used to examine the difficulty of non-adult aging.

To evaluate the full spectrum of facial recognition techniques, six algorithms were used. The algorithms are as follows: FaceVacs, S4F, CohortLDA, LRPCA, LDA and PCA[10, 23, 26, 29, 5, 37]. Both commercial and research based implementations were used to conduct the experiments. Cognitec, the only commercial algorithm used in this work, was chosen due to its strength in nearly all scenarios[10]. The additional five algorithms, were implemented in open-source environments. OpenBiometrics (OpenBR), implemented the S4F algorithm in its evaluation toolkit[23]. Colorado State University implemented two toolkits used in this thesis, the 2011 Baseline Algorithms and the CSU Face Identification Evaluation System[29, 26, 6]. The 2011 Baseline Algorithms is comprised of CohortLDA and LRPCA. The CSU Face Identification Evaluation System implements PCA and LDA which is used in this work. S4F, LRPCA, and CohortLDA are considered more modern techniques. LRPCA uses local regions to extract additional information about the face. LRPCA local regions make it more robust to the orientation of the face across multiple images. CohortLDA is designed around handling illumination in a robust manner by incorporating the I space. Finally S4F was originally designed for the aging space by using spectral sampling. LDA and PCA are both *traditional* algorithms and are sensitive to many changes, whether this be lighting, background, focal length etc.

Each system used in this work was trained using a predefined set used in their papers. We did this to be able to compare these algorithms against other works which have used the same algorithm and training data. OpenBR uses the MEDS dataset for its S4F algorithm. FERET is used in the CSU Face

Identification Evaluation System for LDA and PCA. Finally The Good, The Bad and The Ugly was used in CSU's 2011 Baseline Algorithms(LRPCA and CohortLDA).

The creation of the ITWCC dataset was created through the capture of *in the wild* data. ITWCC contains temporally distanced data throughout the lives of the subjects. Images in ITWCC range from adolescent years into adulthood for each subject. The images in this dataset were collected by using open Internet sources and contains a range of quality, illumination and pose. ITWCC is comprised of 301 subjects and 1705 images. The subjects ages within this dataset range from 5 months to 32 years. The dataset contains 872 female subjects and 846 male images. The average age of all images is 13.418 years with a standard deviation of 3.432 years. The average age of the subjects for acquisition into this dataset is 10.2119 years with a standard deviation of 3.986 years; furthermore, the average age of final capture is 16.265 years with a standard deviation of 4.467 years. Figure 3.7 shows a subset of the ITWCC dataset.

While ITWCC contains a vast amount of non-adult data, it is not large enough to carry out some experimental scenarios. To conduct these scenarios, ITWCC was augmented with The Labeled Faces in the Wild(LFW) Dataset as well as the CASIA twins dataset[16, 32]. LFW was chosen as the acquisition of the photos is similar to how ITWCC was created; however, this dataset does not contain any adolescent data. LFW's diversity is very similar to ITWCC. CASIA twins was chosen for the inclusion of child data, and the partially unconstrained image collection.

Four scenarios were conducted in this thesis to begin to understand the impact of non-adult data. The first scenario is an access control scenario, which includes the ITWCC dataset alone. This scenario matches all 1705 images against all other images to determine the identity of each match. This scenario contains 2,907,025 matches that are evaluated. The next scenario is a photo-tagging scenario, in which ITWCC was used to generate a temporally distanced gallery and probe. This scenario is generated by taking the first photo of each image and placing that into the gallery and placing the remaining into the probe set. This resulted in a gallery of 301 and a probe of 1404. This scenario contains 422,604 matches. The following scenario is an augmented photo-tagging scenario in which additional non-ideal images(LFW



Figure 3.7: In The Wild Child Celebrity Dataset

and CASIA twins) were used as well. Similarly this set uses the previous gallery and probe; however, both the LFW and CASIA twins set was included in the gallery. This resulted in a gallery of 14,768 and a probe of 1,404. By augmenting the CASIA Twins and LFW dataset the number of matches in the augmented photo-tagging scenario climbs to 20,734,272. The final scenario attempts to determine which growth periods pose the greatest challenge for the selected facial recognition algorithms. This scenario breaks down the images into three year age spans and individual sub experiments are conducted for all matches. Due to the size of the ITWCC dataset, some of these sub-experiments did not contain enough information to be conclusive and were omitted from the results.

The age span experiment was separated into age spans of three. Often large changes in adolescence would occur in these age gaps. As shown in figure 2.4 there are large changes in specific ages[35]. For example, from birth to approximately age three there is a large change in eye shape and from ages six to nine there are less prominent changes. The temporal information in these features will effect the performance recognition. The age span scenario will evaluate the different features of the adolescent face which contains aging data to determine the most difficult ages and features for temporally distanced facial recognition.

Each biometric system in this thesis preprocess the images into a form which it can extract the best information from. To achieve this, all the systems use the eye coordinates from each image to extract the face and register it into a standardized format. Cognitec, OpenBR, and CSU's 2011 baselines each use a automated eye detection algorithm to acquire the eye locations. CSU's 2006 Biometric Toolkit require eye coordinates to be provides and where completed by hand in this work. If one of the eye detection algorithms can not find the eyes in an image the image is considered *failure to enroll* and is not included into the matching. PCA and LDA do not have any failure to enroll errors. These systems all continue to process the images after the eye coordinates and the images are registered. An example of a geometrically registered image is shown in figure 3.8. OpenBR implemented a illumination reduction technique proposed by Tan and Triggs[34]. The steps for this method are as follows: Gaussian blur, a difference of Gaussians, gamma correction and finally a contrast equalization. LRPCA uses a registration



Figure 3.8: Normalized ITWCC images

process which requires very strict coordinate locations. LRPCA's registration allows it for it crop out the fourteen sections used in the algorithm. Each of the fourteen sections are then sampled and a self-quotient normalization is used on each individual section. The final step is a zero-mean and stand deviation normalization. CohortLDA uses a logarithmic transformation on the red color space to reduce noise. CohortLDA also uses a zero-mean and standard deviation normalization on both the red and I color spaces to complete its preprocessing. LDA and PCA algorithms require the format of the images to be in PGM greyscale. CSU's 2006 biometric evaluation kit starts by converting the 256 greyscale integers and converts them to float values. It then registers the images and runs a histogram equalization. The final step of the tool uses a zero-mean and standard deviation normalization.

CHAPTER 4: EXPERIMENTAL DESIGN

Four scenarios were used in this work understand the difficulty of temporally displaced data. The scenarios were developed to replicate real world scenarios where this type of data would be often used. As previously discussed four biometric systems were used to detect, preprocess, match and finally evaluate the images. All of the matching information was provided, in this work, to OpenBR's Face Evaluation toolkit[29]. This toolkit evaluates the matching information and then plots the information in a standardized format. Four systems are implemented in this work, to explore the challenges of adolescent aging for facial recognition. Within these four systems, 6 algorithms are evaluated. The algorithms are as follows: FaceVacs, S4F, CohortLDA, LRPCA, LDA and PCA[10, 23, 26, 29, 5, 37].

As previously discussed, by leveraging the In The Wild Child Celebrity dataset a understanding of the non-adult problem can be formed. FaceVacs, S4F, LRPCA, CohortLDA, LDA and PCA are used to establish a general baseline across a spectrum of algorithm types. Cognitec, the only commercial algorithm used, shows the industry standard for this problem. Three algorithms, S4F, CohortLDA, and LRPCA, used in this experiment are considered *modern baselines*. Finally two *traditional* algorithms, LDA and PCA , are used to show a often poor performing system. OpenBR provides an open-source biometric evaluation suite which will be used for this work. Each experiment is separated into a gallery and probe set to answer questions about non-adult aging. Each algorithm is used for every experiment to show the spectrum of techniques.

The first experimental scenario designed for this work mimics an access control scenario. The purpose of the All-to-All Verification experiment is to determine how effective face verification performs when matching between temporally displaced non-ideal images. This experiment compares all images within the ITWCC dataset against all other images. As shown in figure 4.9, images of the same individual are matched against the same individual and all others. The Access Control Scenario was conducted to understand how effective or ineffective the selected algorithms perform for matching adolescent faces. The entire ITWCC dataset was used in this scenario to generate 2,905,320 matches, with 10,652 genuine matches and 2,894,688 matches. Additionally, images that matched to themselves were masked from

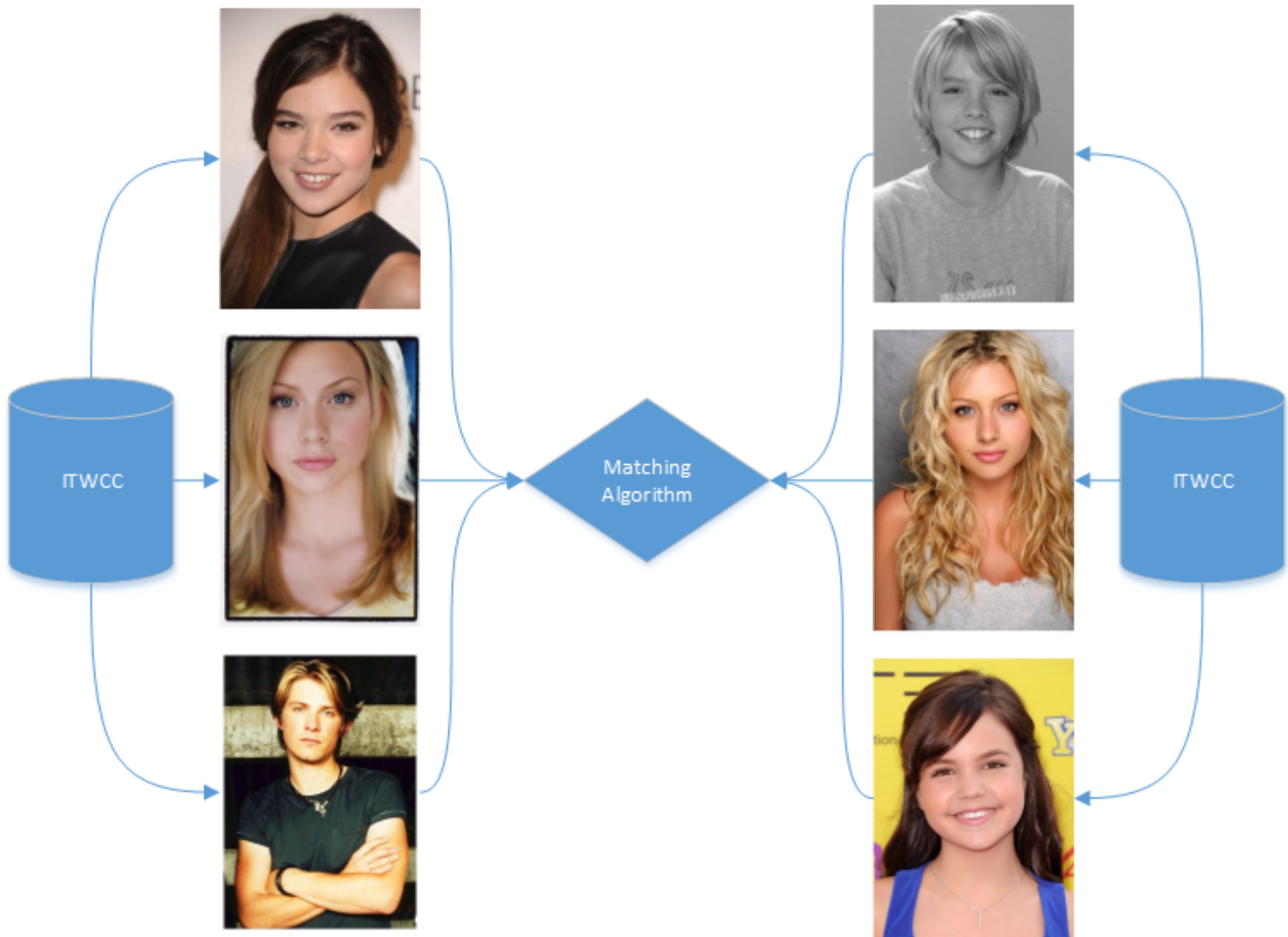


Figure 4.9: Access Control Design

the results. The total number of matches is calculated by: $(\text{Probe Images} * \text{Gallery Images}) - \text{Ignored Images} = \text{Total Matches}$.

Images Used By Dataset		
Dataset Name	Gallery Size	Probe Size
ITWCC	1705	1705

Table 4.1: All-to-All Verification Experiment Dataset Usage

The second experiment is designed to emulate a photo-tagging across time scenario. The Young-to-Older Identification experiment's objective is to determine the effectiveness of the selected algorithms against monotonically increasing ages. This experiment attempts to setup a scenario in which an end-user of a photo tagging tool, such as Facebook, Picasa, etc., would begin adding images over a span of

Matching Comparisons			
Genuine Matches	Impostor Matches	Ignored Matches	Total Matches
10,652	2,894,668	1,705	2,905,320

Table 4.2: All-to-All Verification Matching Comparison

time[1, 14]. The ITWCC dataset is used in this experiment similarly to the first experiment; however, only the first image of each person is used for the gallery and all other images for the individual are used in the probe. Shown in figure 4.10, the youngest image is matched to all of its elder images. The average age of the gallery was 11.81 with a standard deviation of 4.23. The remaining images were then placed in the probe set. The average age of the probe set is 11.61 with a standard deviation of 3.92. These averages represent the data as a whole and does not reflect the individual age differences between each subject.

Images Used By Dataset		
Dataset Name	Gallery Size	Probe Size
ITWCC	301	1404

Table 4.3: Young-to-Older Identification Experiment Dataset Usage

Matching Comparisons			
Genuine Matches	Impostor Matches	Ignored Matches	Total Matches
1404	421,200	0	422,904

Table 4.4: Young-to-Older Identification Matching Comparison

The augmented Young-to-Old identification experiment further extends the last experiment by increasing the gallery size. The gallery is augmented with both the CASIA Twins and the Labeled Faces in the Wild datasets[32, 16]. By expanding the gallery with both of these datasets, the scenario will be closer to a real world situation in which a user would upload additional data to match against. This scenario is expected to be much more challenging for identification across non-ideal images. All of the CASIA Twins and LFW are added to the gallery, which results in a gallery size of 14,764 images versus 301 from the previous Young-to-Old Scenario. LFW does not contain any adolescent data; however, the data is captured in a similar manner to ITWCC. The CASIA Twins dataset does contain adolescent data,

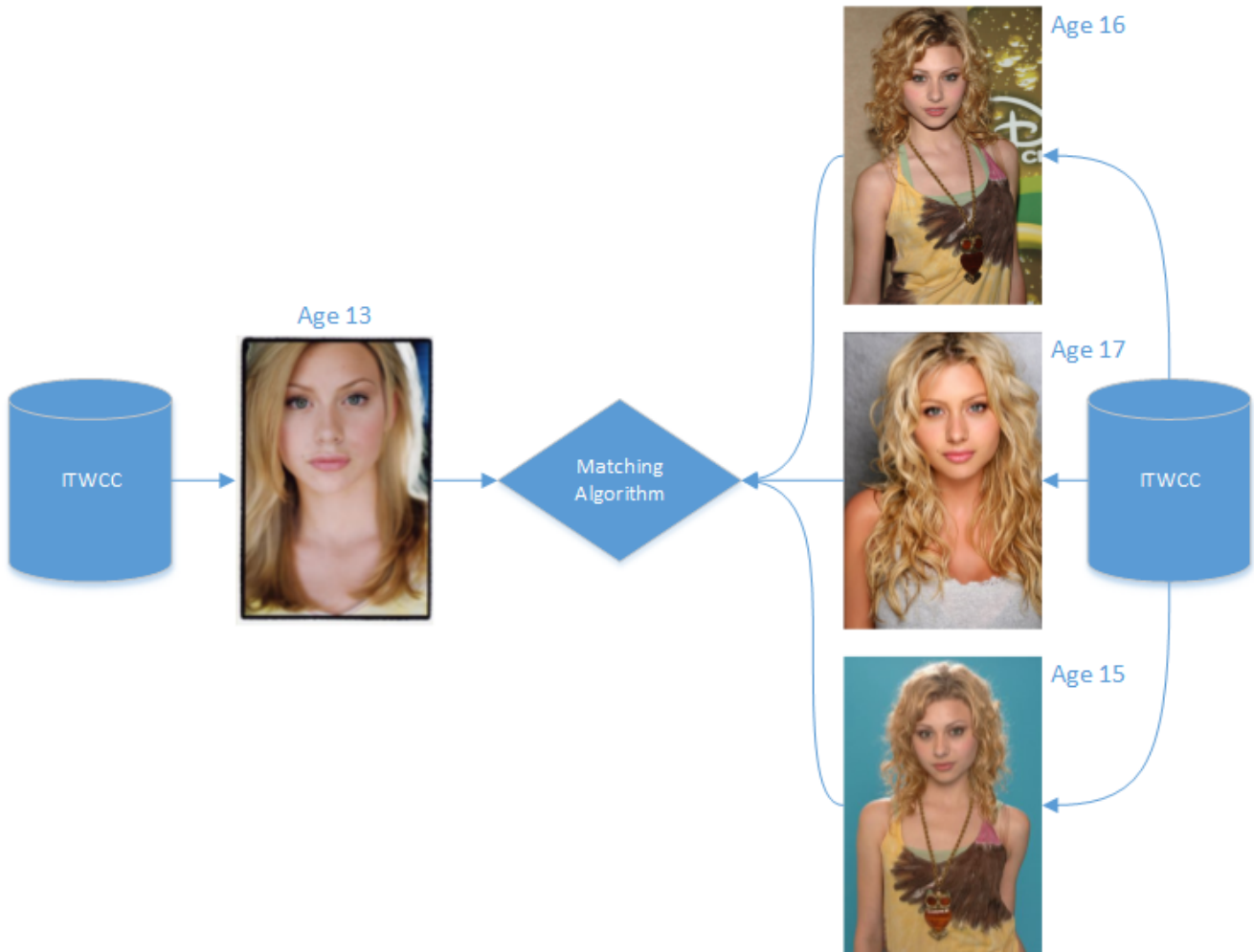


Figure 4.10: Photo Tagging Design

but it is captured in a slightly less varied manor.

Images Used By Dataset		
Dataset Name	Gallery Size	Probe Size
ITWCC	301	1404
LFW	13,233	0
Twins	1,234	0

Table 4.5: Augmented Young-to-Older Identification Experiment Dataset Usage

Matching Comparisons			
Genuine Matches	Imposter Matches	Ignored Matches	Total Matches
1404	20,732,868	0	20,734,272

Table 4.6: Augmented Young-to-Older Identification Matching Comparison

The final experiment, Growth Period Verification, attempts to explore the difficulty of each temporal period of adolescence. The scenario is designed around growth periods which were deemed challenging due to large facial changes. In Karen Taylors work, she describes the facial deformations of an adolescent male, shown in figure 2.4[35]. Due to the non-uniform facial growth patterns, differing changes may be more impactful on recognition algorithms than others. By matching difficult growth period in adolescence we can gain insight on what ages pose the most challenge for modern facial recognition techniques. This was accomplished by separating each images into sets of three years i.e. 0-2, 3-5, 6-8, . . . ,18+. Each age set was matched to each other age set as shown in figure 4.12. Individuals were separated into probe and gallery again by comparing youngest photos in the gallery to the eldest photos in the probe. The ITWCC dataset was used solely for the Growth Period Verification tests. However, due to the size of the dataset, some of the experiments were excluded from the results. The size of the dataset did not allow for certain age to age experiments to have significant enough information to derive any conclusions. Any age-to-age scenario with less than 1000 total matches between all images are removed from the results. The number of matches is shown in figure 4.8. Unlike other experiments in this paper, the arrangement of probe and gallery was chosen based upon the age ranges and not just the youngest photo. Images within the gallery are acquired from the lowest age set and the probe captured from the eldest age set.

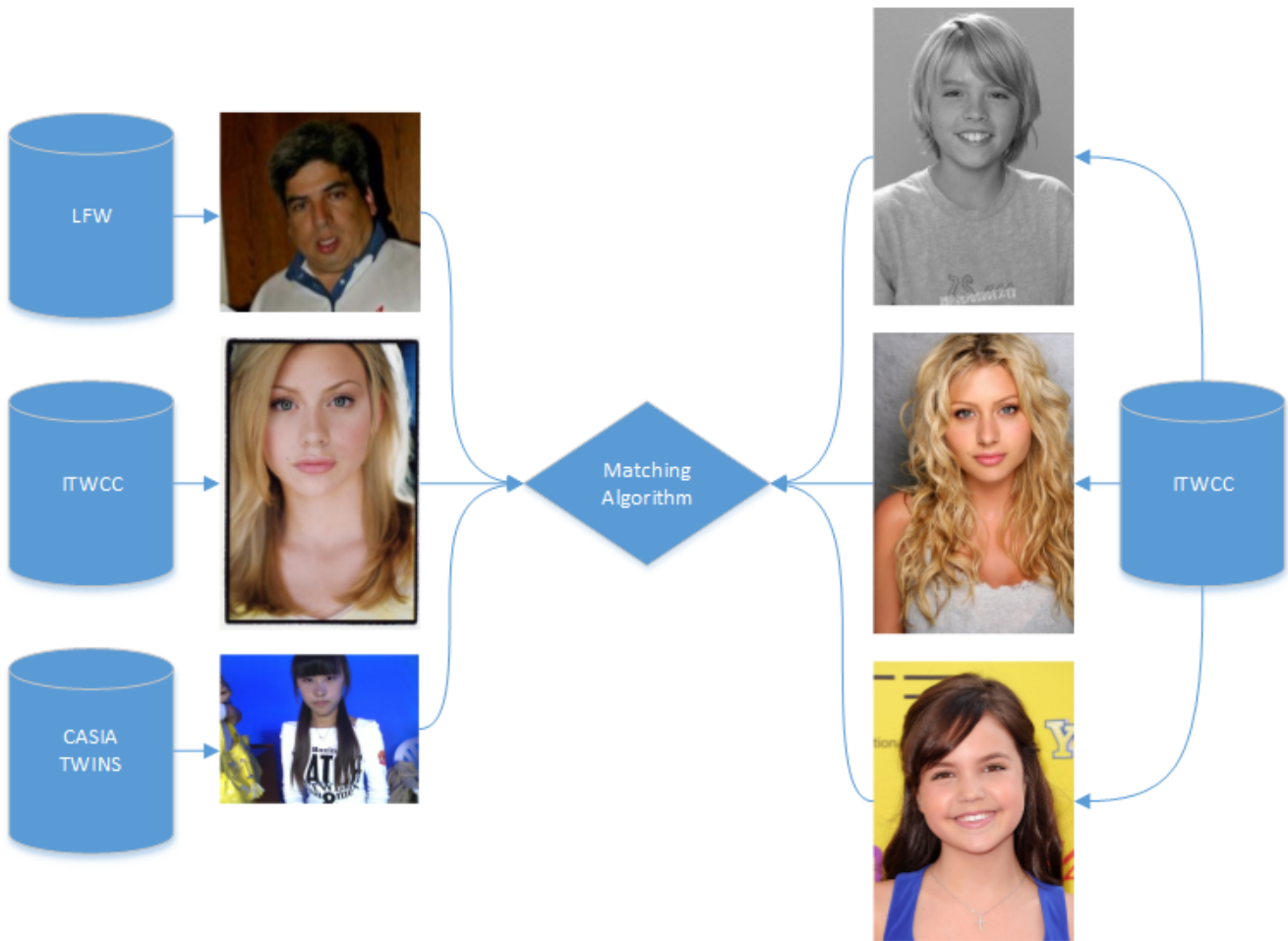


Figure 4.11: Augmented Photo Tagging Design

Images Used within ITWCC		
Age Range	Gallery Size	Probe Size
0-2 to 3-5	6	5
0-2 to 6-8	4	5
0-2 to 9-11	3	6
0-2 to 12-14	3	6
0-2 to 15-17	6	13
0-2 to 18+	3	4
3-5 to 6-8	30	33
3-5 to 9-11	27	58
3-5 to 12-14	33	70
3-5 to 15-17	19	32
3-5 to 18+	9	17
6-8 to 9-11	73	118
6-8 to 12-14	57	94
6-8 to 15-17	29	42
6-8 to 18+	14	16
9-11 to 12-14	222	265
9-11 to 15-17	84	136
9-11 to 18+	56	48
12-14 to 15-17	281	319
12-14 to 18+	120	129
15-17 to 18+	207	179

Table 4.7: Growth Period Verification Experiment Dataset Usage

Matching Comparisons				
Age Range	Genuine Matches	Imposter Matches	Ignored Matches	Total Matches
0-2 to 3-5	9	21	0	30
0-2 to 6-8	5	15	0	20
0-2 to 9-11	6	12	0	18
0-2 to 12-14	6	12	0	18
0-2 to 15-17	13	65	0	78
0-2 to 18+	4	8	0	12
3-5 to 6-8	43	947	0	990
3-5 to 9-11	73	1493	0	1566
3-5 to 12-14	88	2222	0	2310
3-5 to 15-17	36	572	0	608
3-5 to 18+	17	136	0	153
6-8 to 9-11	213	8401	0	8614
6-8 to 12-14	160	5198	0	5358
6-8 to 15-17	75	1143	0	1218
6-8 to 18+	23	201	0	224
9-11 to 12-14	613	58217	0	58830
9-11 to 15-17	225	11199	0	11424
9-11 to 18+	84	2604	0	2688
12-14 to 15-17	805	88834	0	89639
12-14 to 18+	247	15233	0	89639
15-17 to 18+	418	36635	0	37053

Table 4.8: Growth Period Verification Matching Comparison

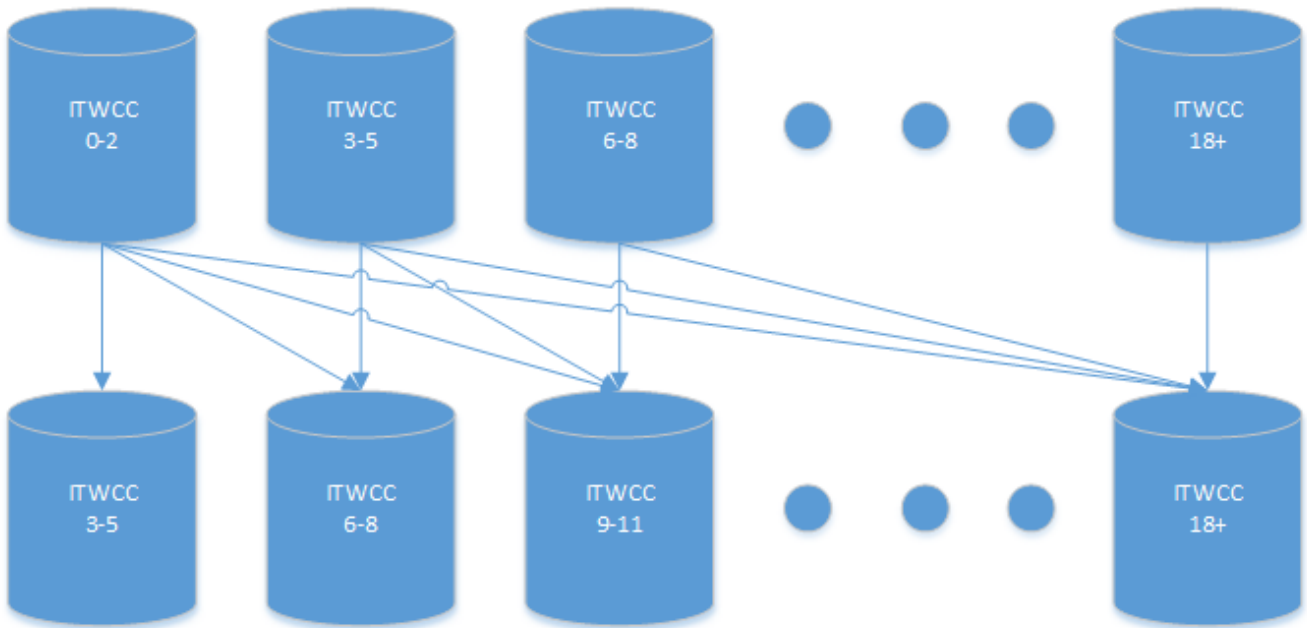


Figure 4.12: Growth Period Verification Design

CHAPTER 5: RESULTS

This work provides four scenarios to begin to understand the challenges of facial recognition on temporally distanced non-adult faces. The four scenarios are as follows: Access control or All-to-All Verification, Photo-tagging or Young-to-Old Identification, augmented photo-tagging or Augmented Young-to-Old Identification and a growth period verification. Due to the size of ITWCC, some experiments were not provided as they did not show any substantial information. I have provided receiver operating characteristic, detection error trade-off, Score Histograms and other graphs to help in recognizing patterns where age may become a degrading factor for facial recognition. All graphs were generated using the OpenBR Biometric Toolkit[23].

Each graph provided in this work shows a particular comparison to determine strengths and weaknesses in each algorithm. True accept rates, false accept rates, true rejection rates and false rejection rates provide insight on the number of correct and false positives for both errors and correct acceptances. Receiver operating characteristic(ROC) shows the accuracy of a biometric system, by comparing the false acceptance rate vs. the verification rate. For this particular characteristic, higher values indicate higher accuracy. Detection Error Trade-off(DET), plots the error rates of each algorithm. The false reject rate is plotted against the false accept rate to determine the error rates for each algorithm. Score histograms plot the frequency of verification scores. Values are listed from 0 - 1 and both genuine and imposters are shown on this chart. An ideal plot would have genuine matches nearing 1 while imposters nearing 0 with a clear division between them. Cumulative match characteristic (CMC) shows the accuracy of a closed-set identification process. During the identification process all images are scored and ranked in an ascending order based on their match score. The CMC graph shows the percentage that the genuine user is in a rank less than or equal to the current rank being plotted.

Shown in figure 5.9, the true accept rates of the All-to-All Verification, the Young-to-Old Identification, and the Augmented Young-to-Old Identification provide a way to estimate the performance of each of the algorithms in this work. This figure gives an estimation of the accuracy of each algorithm, in particular, how often an impostor gains access to the system vs. how often a true user is accepted

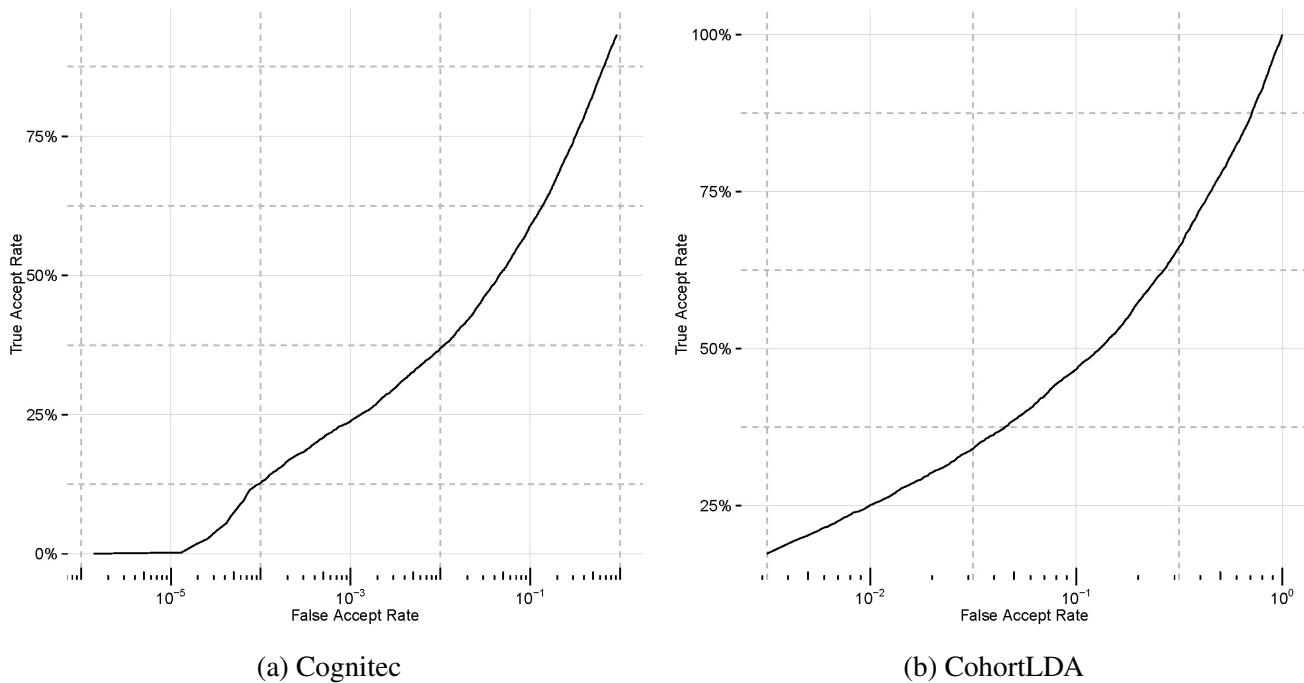


Figure 5.13: Best and Worst ROC curves for All-to-All Verification

when the system is tuned to have a false acceptance rate of 1%. The Access control scenario shows that Cognitec’s system performed significantly better than any of the others, at 37% true acceptance rate. Figure 5.13 shows the best, Cognitec, and worst performing, CohortLDA, algorithms for access control. Once again in the Young-to-old Identification experiment, Cognitec outperforms all other algorithms by a fairly significant margin and again S4F coming up second with 30.3% and 21.5% true acceptance rates respectively.

Another approximation of performance is shown through the score histograms. Figure 6.17 charts the frequency of the verification scores(normalized between 0-1) for All-to-All Verification. A 0 indicates a poor match and a 1 indicated that its very similar. Figure 6.23 shows the Young-to-Old Identification Experiment’s score histogram. Again with both the All-to-All Verification and Young-to-Old Identification experiment, Cognitec and S4F performed better in terms of clearly separating matches. Cognitec’s score frequency shows that impostors are generally towards 0. However, the frequency for genuine users was also distributed closer to 0 with a significant spike in frequency close to 1. CohortLDA, LRPCA,

LDA and PCA all had very similar frequency in respect to similarity with genuine users and impostors.

Detection error trade-off(DET) demonstrates yet another way to show performance of a face recognition system. By plotting the false accept rate against the false reject rate, one can determine how often a individual is miss-classified. shown in figure 6.16. In both the All-to-All verification and Young-to-Old identification, shown in figure 6.22, Cognitec produced less errors than all the other systems. The equal error rate(EER) is determined by finding the intersection of the false accept rate and the false reject rate.

In the All-to-All Verification and Young-to-Old Verification experiments, Cognitec in general showed better performance; however, when the young-to-old identification experiment was augmented the with LFW and CASIA Twins, true accept rate dropped to 0.8% from 30.3%. [16, 32] In figure 5.9 you can see that Cognitec’s ability to correctly identify a genuine user plummeted when noise was introduced in the form of the LFW and CASIA Twins datasets. Shown in figure 6.29, the normalized score histograms for this experiment. While Cognitec often correctly scores impostors, it also incorrectly scored genuine users. For example, in the Young-to-Old Identification experiment, a 50% retrieval rank for Cognitec, shown in figure 6.24, is at approximately rank 27 while, in the Augmented Young-to-Old Identification experiment, at rank 27 the retrieval rate is less than 0.01%.

True Accept Rates			
Algorithm Name	All-to-All Verif.	Young-to-Old Ident.	Aug. Young-to-Old Ident.
Cognitec	37%	30.3%	0.8%
S4F	25%	21.5%	32.1%
CohortLDA	12.1%	8.8%	13.2%
LRPCA	13.5%	10.7%	16.1%
LDA	12.6%	10.8%	19.9%
PCA	15%	12.4%	23.5%

Table 5.9: True Accept Rate at 1% False Accept Rate

The growth period verification experiment was challenging to produce with the small size of the ITWCC dataset. Figure 4.8 shows the matches for this experiment. All tests provided in the results have at least 1000 total matches, as any matches below that would be inconclusive. In terms of true accept

True Accept Rates			
Algorithm Name	All-to-All Verif.	Young-to-Old Ident.	Aug. Young-to-Old Ident.
Cognitec	23.8%	19.1%	0.0%
S4F	5.5%	5.7%	1.7%
CohortLDA	2.7%	2.0%	3.9%
LRPCA	5.9%	4.7%	8%
LDA	4.6%	3.8%	8.5%
PCA	5.4%	4.8%	10.5%

Table 5.10: True Accept Rate at 0.1% False Accept Rate

rates, Cognitec outperformed all other systems for this experiment. Shown below in figure 5.11, are the true accept rates of Cognitec’s system for each age range used in this work. The table is sorted in by the average age difference of each range in descending order. You can see on average that as the average age difference between the age ranges minimizes, the performance increases. Large facial feature changes in the growth and development stages where expected to have a large impact on performance, i.e. at age 4 pronounced chin growth occurs as well as age 12 where the chin line becomes more defined. These two features will look significantly different at these ages; however, the findings did not show this. The only thing difference that seemed to correlate with the results is the age span difference. Due to the inconsistencies in the dataset size for each of these evaluations may not be conclusive; however, with a large enough dataset one could prove the significance of this problem.

True Accept Rates		
Growth Period Range	TAR	Average Age Difference
3-5 to 12-14	27.8%	9
6-8 to 15-17	16.6%	9
9-11 to 18+	19.1%	8
3-5 to 9-11	13.7%	6
6-8 to 12-14	24.5%	6
9-11 to 15-17	26.3%	6
12-14 to 18+	29.6%	5
6-8 to 9-11	30.8%	3
9-11 to 12-14	36%	3
12-14 to 15-17	38.3%	3
15-17 to 18+	45.6%	2

Table 5.11: Cognitec’s FaceVacs True Accept Rate at 1% False Accept Rate

CHAPTER 6: CONCLUSION

This work examined the challenges of facial recognition over temporally distanced non-adult faces with the introduction of the In the Wild Child Celebrity (ITWCC) dataset. Four experimental scenarios were designed to explore the difficulty of non-adult aging by leveraging the ITWCC, CASIA Twins, and LFW datasets [32, 16]. The scenarios are as follows: Access control, Photo-tagging, augmented photo-tagging, and growth period verification. This work compared six algorithms to test the hypothesis: Cognitec’s FaceVacs, OpenBR’s Spectrally Sampled Structural Subspaces Features (S4F), Cohort Linear Discriminant Analysis (CohortLDA), Local Region Principle Component Analysis (LRPCA), Linear Discriminant Analysis (LDA), and Principle Component Analysis (PCA) [10, 23, 26, 29, 5, 37]. Results on this unique, albeit small dataset, shows that aging on non-adults is a challenging task for facial recognition algorithms. The most accurate algorithm for the verification task had a true accept rate of only 37.0% at 1.0% false accept rate. In a similar study conducted by Klare et al., a true accept rate of 96.3% was achieved on adults with 0-1 years of lapse between images [21]. The difficulty of the aging problem is further explored within this work by examining performance over age periods. The growth period verification task matches young images of a particular subject to their older images within a specific age span. The most accurate algorithm had an average true accept rate of only 22.51% at a false accept rate

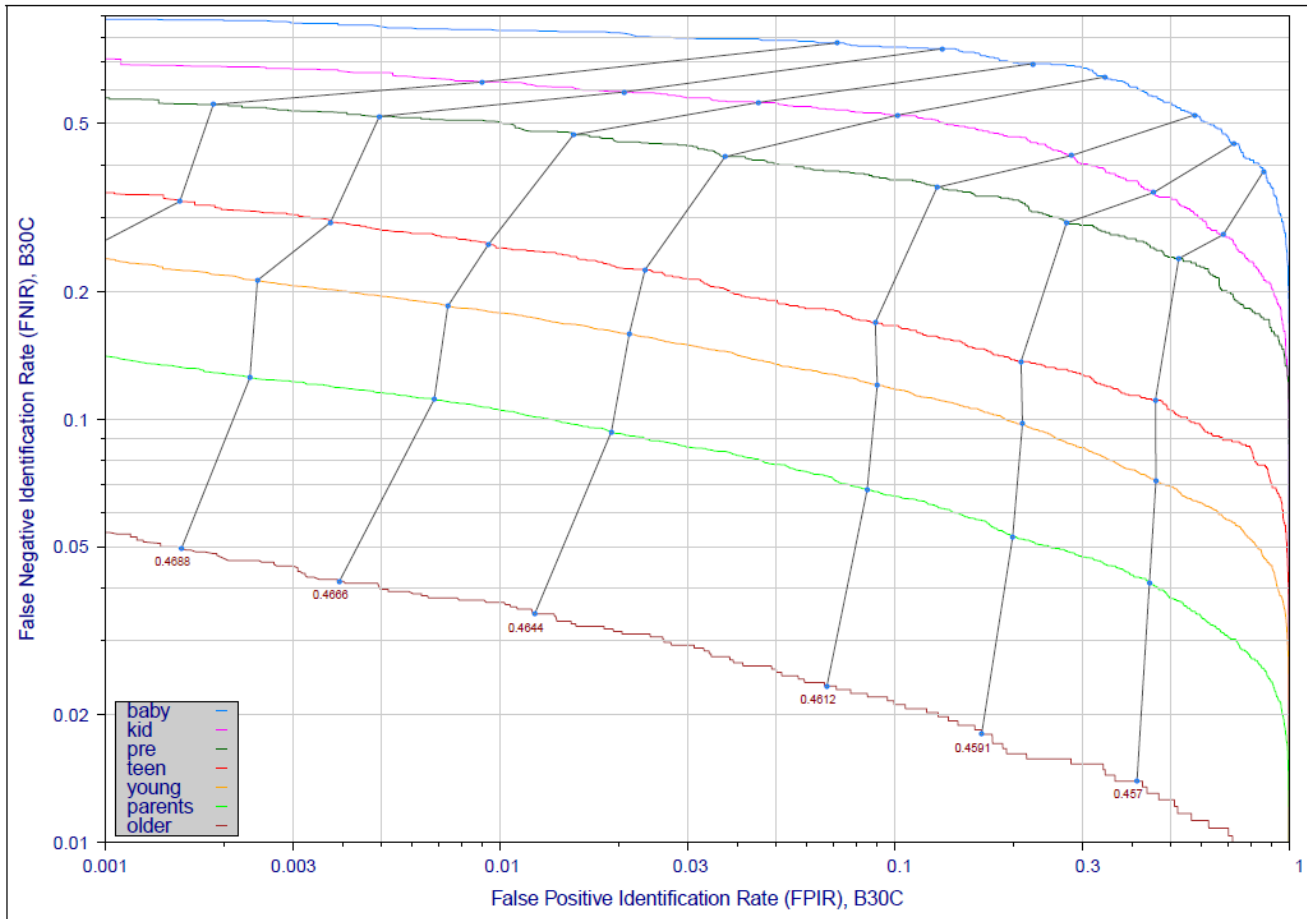


Figure 6.14: NIST Cognitec Results

of 1.0% within the age span of 5 to 10. Again in Klare et al.'s study, a true accept rate of 88.6% was achieved in this age span. The National Institute of Standards and Technology conducted an evaluation of facial recognition algorithms on multiple age groups[28]. Shown in figure 6.14, the study shows the error rates increase as the age group decreased in age. This detection error trade-off curve displays a significant performance reduction across the age ranges, where older is the control in this experiment. This figure displays how often a impostor is allowed into the system in comparison to the rate in which genuine users are rejected. We can conclude that non-adult facial recognition, specifically when across temporally shifted images, is a challenge for facial recognition systems. Additional research into this field will be necessary to uncover the exact causes to this performance degradation.

REFERENCES

- [1] Facebook.
- [2] G. P. A and H. G. J. Factors influencing the accuracy of age estimates of unfamiliar faces. *Perception*, 24:1059–1073, 1995.
- [3] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(12):2037–2041, 2006.
- [4] A. M. Albert, K. Ricanek Jr, and E. Patterson. A review of the literature on the aging adult skull and face: Implications for forensic science research and applications. *Forensic Science International*, 172(1):1–9, 2007.
- [5] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(7):711–720, 1997.
- [6] R. Beveridge, D. Bolme, M. Teixeira, and B. Draper. The csu face identification evaluation system users guide: version 5.0. *Computer Science Department, Colorado State University*, 2(3), 2003.
- [7] D. Bolme, B. Draper, and J. Beveridge. Average of synthetic exact filters. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 2105–2112, June 2009.
- [8] D. S. Bolme. *Elastic bunch graph matching*. PhD thesis, Colorado State University, 2003.
- [9] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.
- [10] Cognitec. Facevac software developer kit, 2014.
- [11] S. R. Coleman and R. Grover. The anatomy of the aging face: volume loss and changes in 3-dimensional topography. *Aesthetic surgery journal*, 26(1 suppl):S4–S9, 2006.

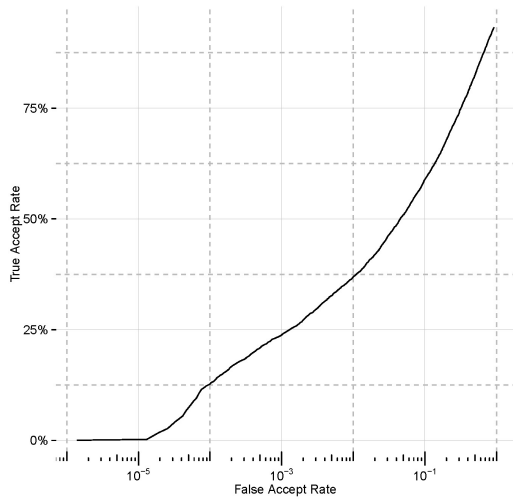
- [12] L. G. Farkas, J. C. Posnick, T. M. Hreczko, and G. E. Pron. Growth patterns of the nasolabial region: a morphometric study. *The Cleft Palate-Craniofacial Journal*, 29(4):318–324, 1992.
- [13] G. H. Givens, J. R. Beveridge, B. A. Draper, and D. Bolme. Using a generalized linear mixed model to study the configuration space of a pca+ lda human face recognition algorithm. In *Articulated Motion and Deformable Objects*, pages 1–11. Springer, 2004.
- [14] Google. Picasa, 2014.
- [15] N. N. H. Schneiderman and M. Sipe. Pittpatt. 2010.
- [16] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical report, Technical Report 07-49, University of Massachusetts, Amherst, 2007.
- [17] A. Jain, L. Hong, and S. Pankanti. Biometric identification. *Commun. ACM*, 43(2):90–98, Feb. 2000.
- [18] A. K. Jain, S. C. Dass, K. Nandakumar, and K. N. Soft biometric traits for personal recognition systems. In *Proceedings of International Conference on Biometric Authentication, Hong Kong*, pages 731–738, 2004.
- [19] K. R. Jr. and B. Barbour. What are soft biometrics and how can they be used? *IEEE Computer*, 44(9):106–108, 2011.
- [20] B. Klare. Spectrally sampled structural subspace features (4sf). *Michigan State University Technical Report, MSU-CSE-11-16*, 2011.
- [21] B. Klare and A. K. Jain. Face recognition across time lapse: On learning feature subspaces. In *Biometrics (IJCB), 2011 International Joint Conference on*, pages 1–8. IEEE, 2011.

- [22] B. F. Klare, M. J. Burge, J. C. Klontz, R. W. Vorder Bruegge, and A. K. Jain. Face recognition performance: Role of demographic information. *Information Forensics and Security, IEEE Transactions on*, 7(6):1789–1801, 2012.
- [23] J. C. Klontz, B. F. Klare, S. Klum, A. K. Jain, and M. J. Burge. Open source biometric recognition. In *Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on*, pages 1–8. IEEE, 2013.
- [24] A. Lanitis. Fgnet aging data base.
- [25] D. G. Lowe. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, volume 2, pages 1150–1157, 1999.
- [26] Y. M. Lui, D. Bolme, P. J. Phillips, J. R. Beveridge, and B. A. Draper. Preliminary studies on the good, the bad, and the ugly face recognition challenge problem. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*, pages 9–16. IEEE, 2012.
- [27] L. Mark, J. Pittenger, and H. Hines. Winkling and head shape as coordinated sources of age level information. *Perception and Psychophysics*, 27:117–124, 1980.
- [28] M. Ngan and P. Grother. Face recognition vendor test (frvt) performance of automated age estimation algorithms. *NIST Interagency Report 7995*, mar 2014.
- [29] P. J. Phillips, J. R. Beveridge, B. A. Draper, G. Givens, A. J. O’Toole, D. S. Bolme, J. Dunlop, Y. M. Lui, H. Sahibzada, and S. Weimer. An introduction to the good, the bad, & the ugly face recognition challenge problem. In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*, pages 346–353. IEEE, 2011.
- [30] P. J. Phillips, P. Grother, R. Micheals, D. M. Blackburn, E. Tabassi, and M. Bone. Face recognition vendor test 2002. In *Analysis and Modeling of Faces and Gestures, 2003. AMFG 2003. IEEE International Workshop on*, page 44. IEEE, 2003.

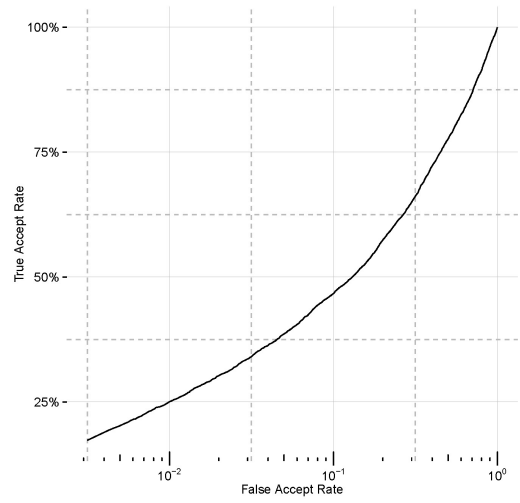
- [31] K. Ricanek, A. Sethuram, E. K. Patterson, A. M. Albert, and E. J. Boone. Craniofacial aging. *Wiley Handbook of Science and Technology for Homeland Security*, 2009.
- [32] Z. Sun, A. A. Paulino, J. Feng, Z. Chai, T. Tan, and A. K. Jain. A study of multibiometric traits of identical twins. In *SPIE Defense, Security, and Sensing*, pages 76670T–76670T. International Society for Optics and Photonics, 2010.
- [33] H. S. D. Taister, Michael A. Comments on facial aging in law enforcement investigation. *Forensic Science Communications*, 2(2), 2000.
- [34] X. Tan and B. Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *Image Processing, IEEE Transactions on*, 19(6):1635–1650, 2010.
- [35] K. T. Taylor. *Forensic art and illustration*. CRC Press, 2010.
- [36] M. L. Teixeira. *The bayesian intrapersonal/extrapersonal classifier*. PhD thesis, Colorado State University, 2003.
- [37] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of cognitive neuroscience*, 3(1):71–86, 1991.
- [38] M. Turk and A. Pentland. Face recognition using eigenfaces. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on*, pages 586–591, 1991.
- [39] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–511. IEEE, 2001.
- [40] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *Acm Computing Surveys (CSUR)*, 35(4):399–458, 2003.

APPENDICES A

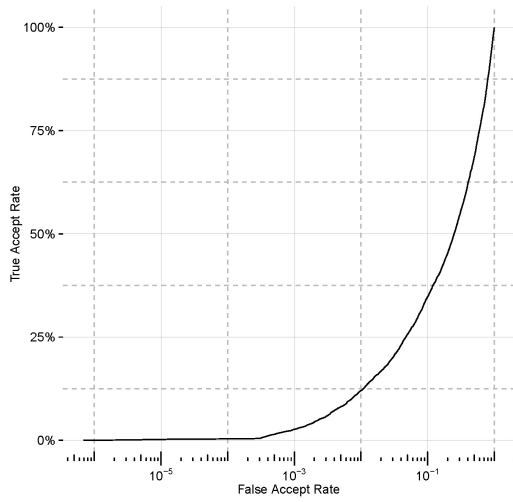
All-to-All Verification Results



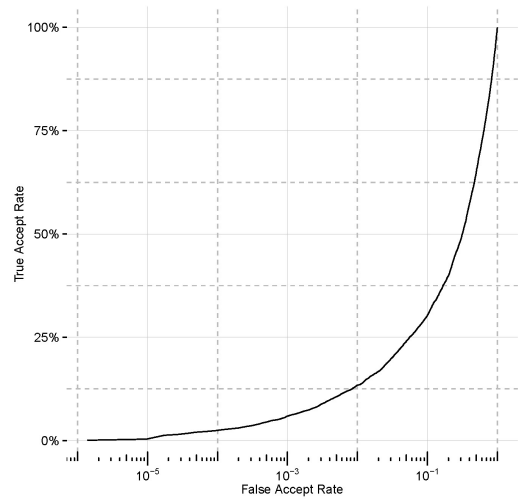
(a) Cognitec



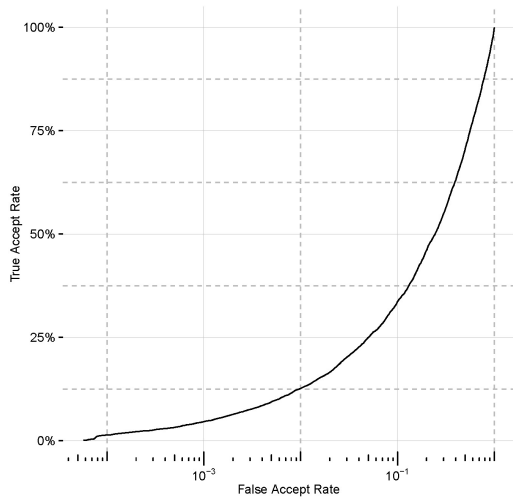
(b) S4F



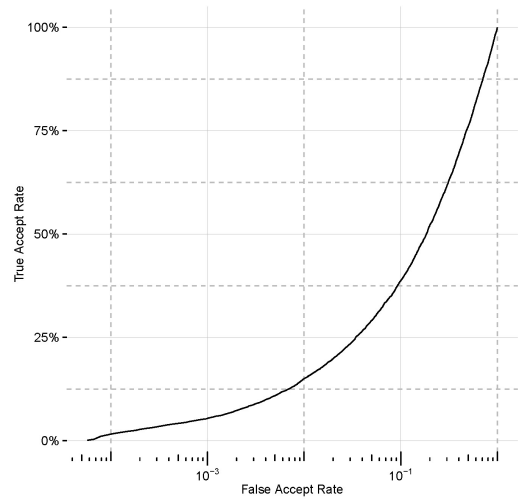
(c) CohortLDA



(d) LRPCA

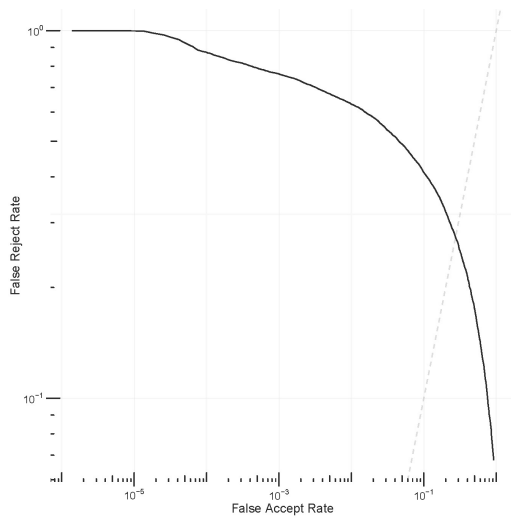


(e) LDA

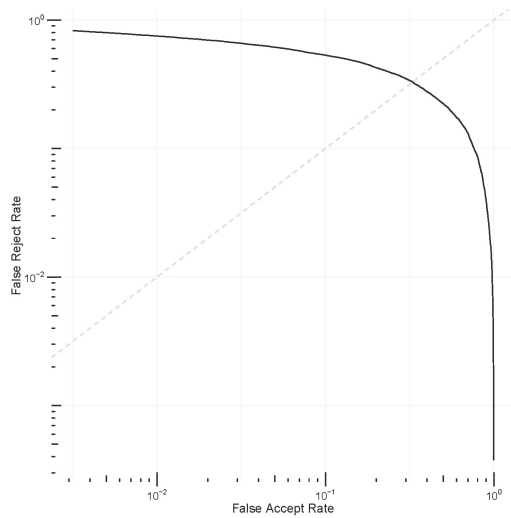


(f) PCA

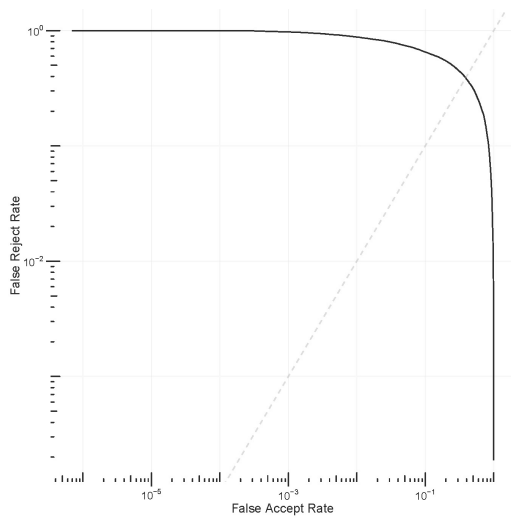
Figure 6.15: Child Celebrity Dataset - ROC



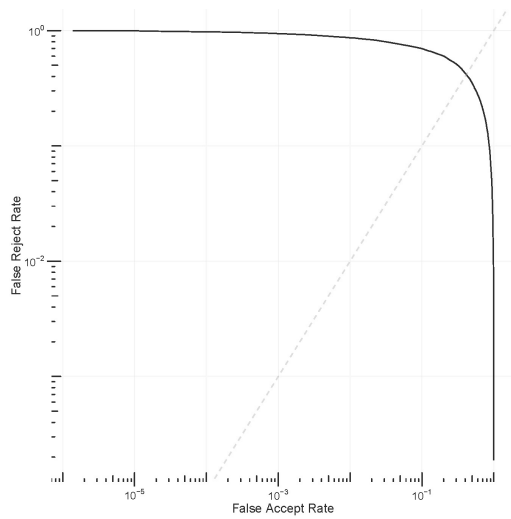
(a) Cognittec



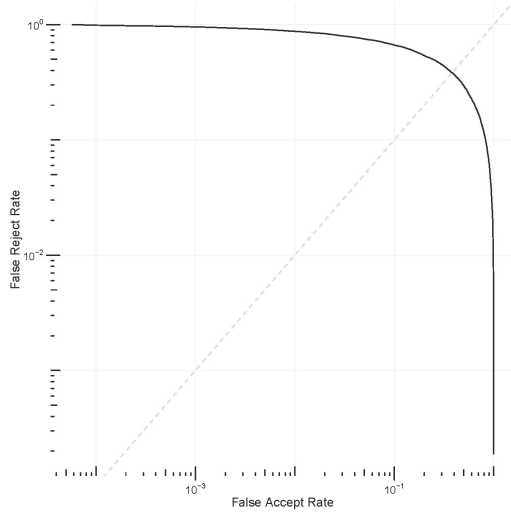
(b) S4F



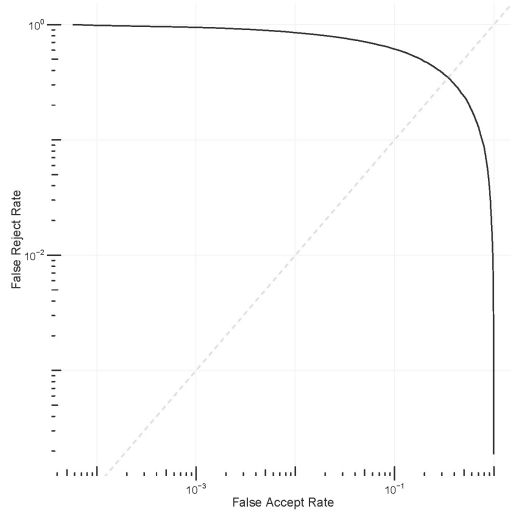
(c) CohortLDA



(d) LRPCA

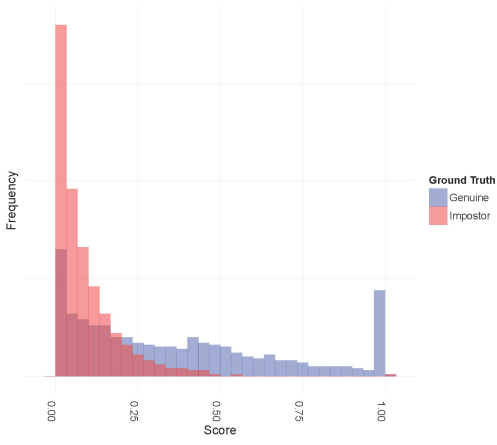


(e) LDA

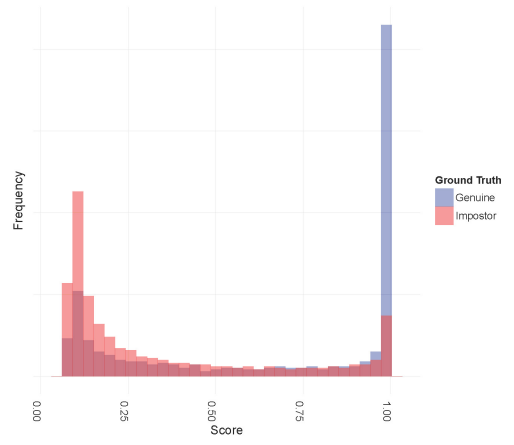


(f) PCA

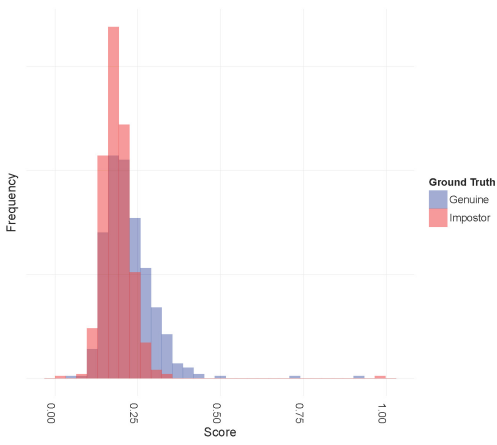
Figure 6.16: Child Celebrity Dataset - DET



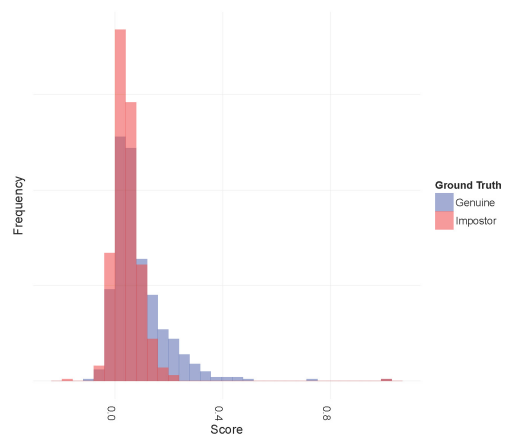
(a) Cognitec



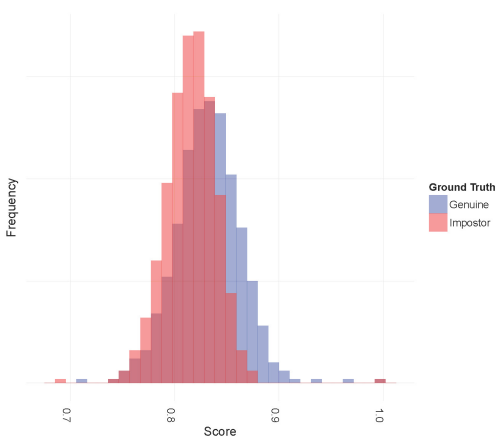
(b) S4F



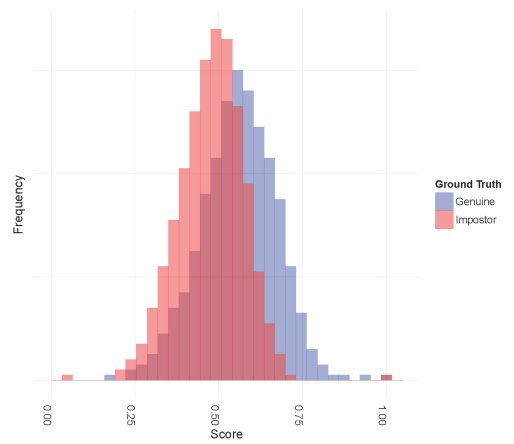
(c) CohortLDA



(d) LRPCA

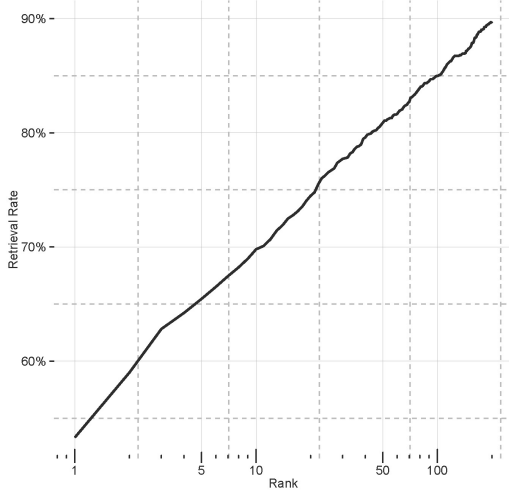


(e) LDA

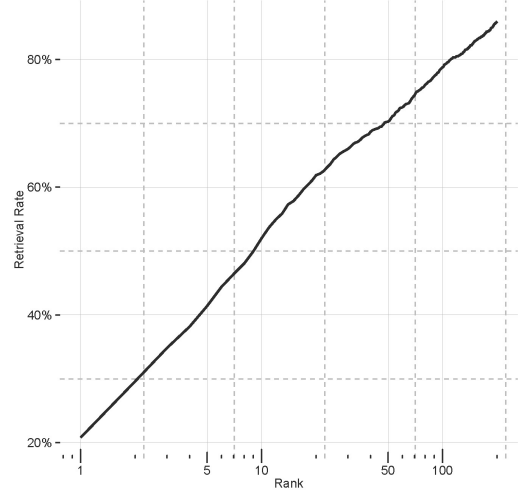


(f) PCA

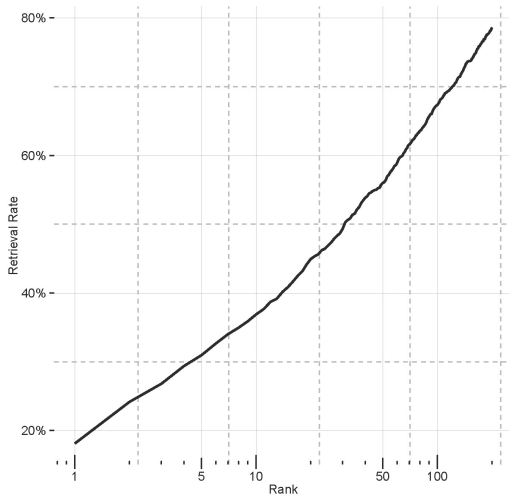
Figure 6.17: Child Celebrity Dataset - Score Histogram



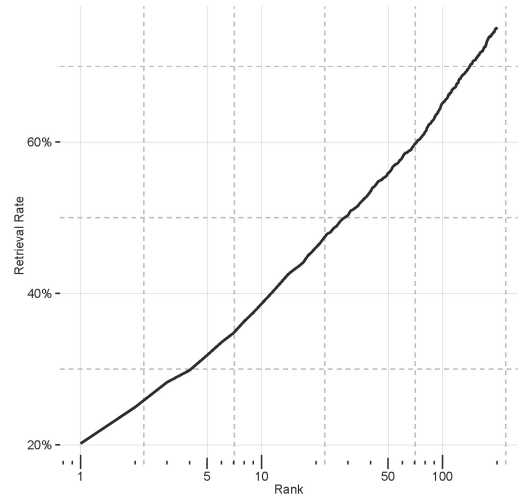
(a) Cognitec



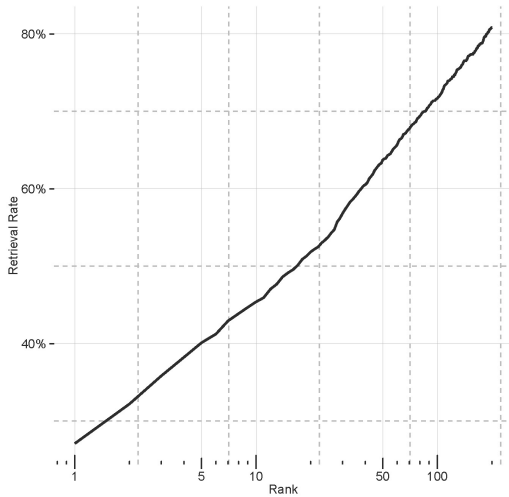
(b) S4F



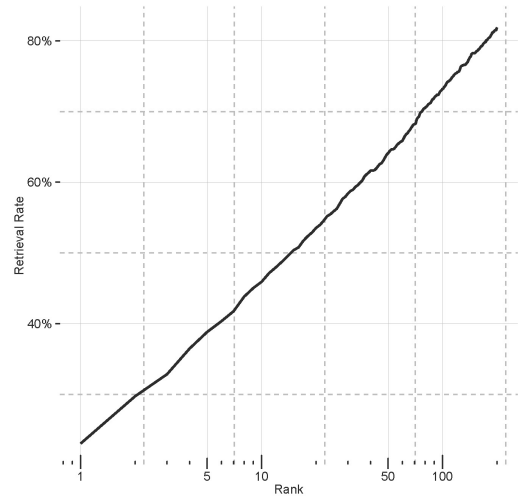
(c) CohortLDA



(d) LRPCA

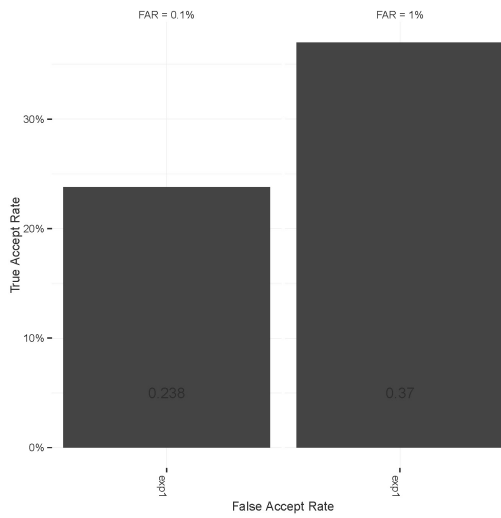


(e) LDA

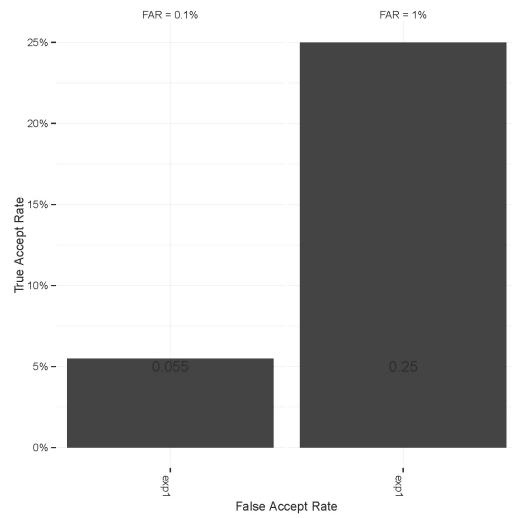


(f) PCA

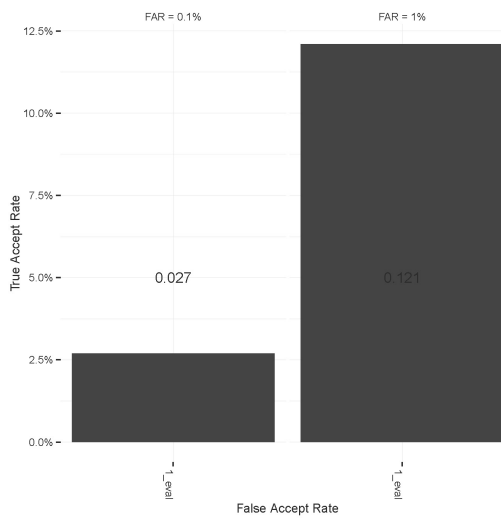
Figure 6.18: Child Celebrity Dataset - CMC



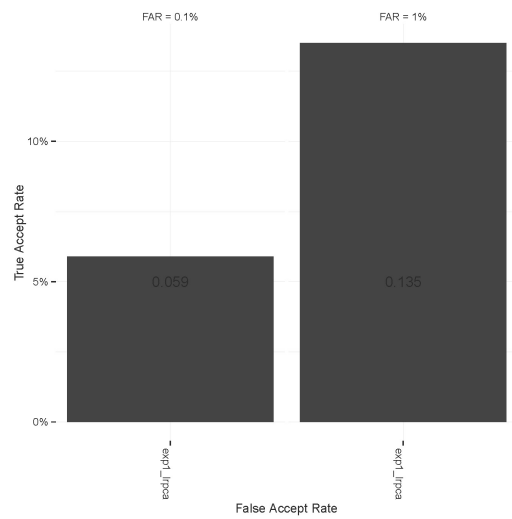
(a) Cognitec



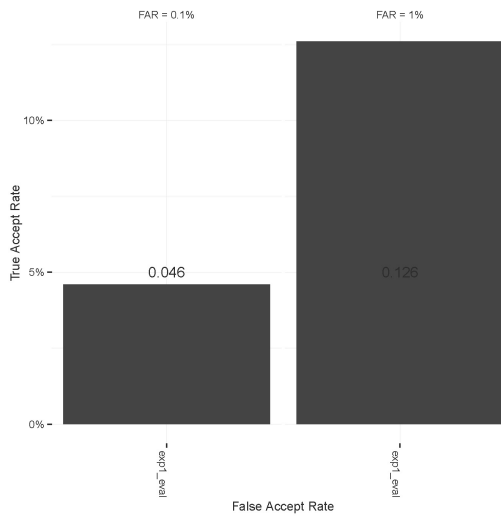
(b) S4F



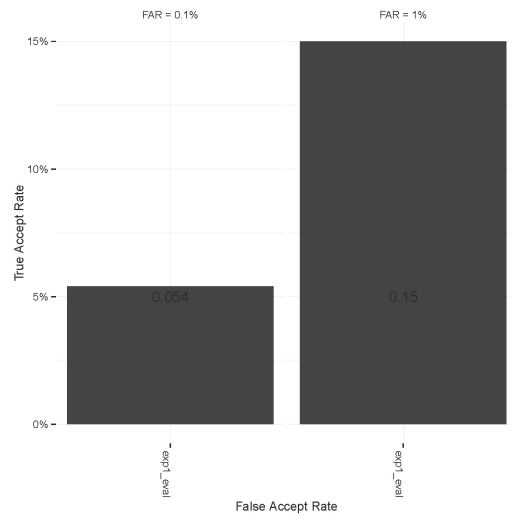
(c) CohortLDA



(d) LRPCA

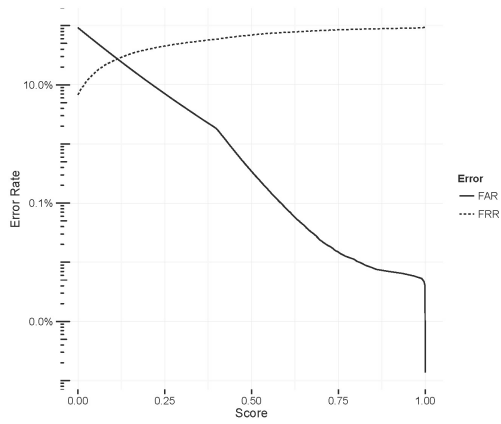


(e) LDA

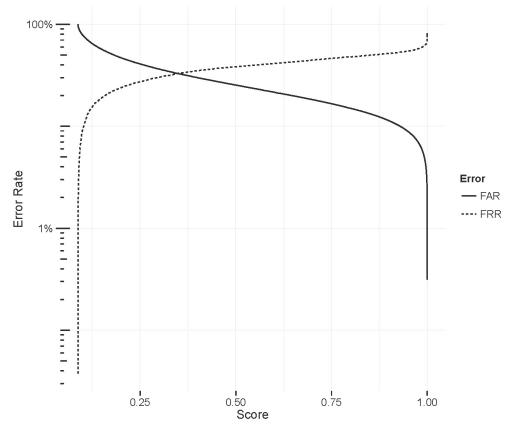


(f) PCA

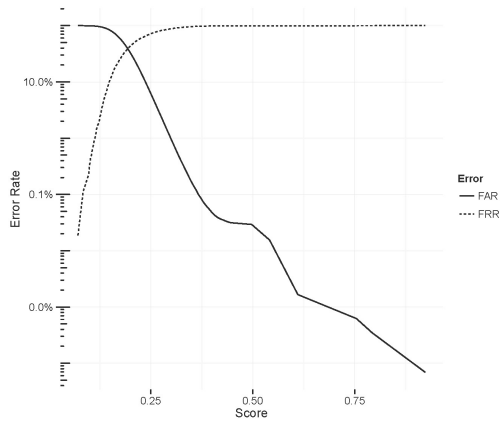
Figure 6.19: Child Celebrity Dataset - ROC Scores



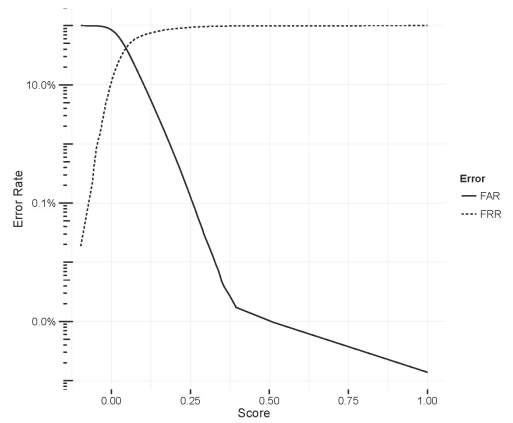
(a) Cognitec



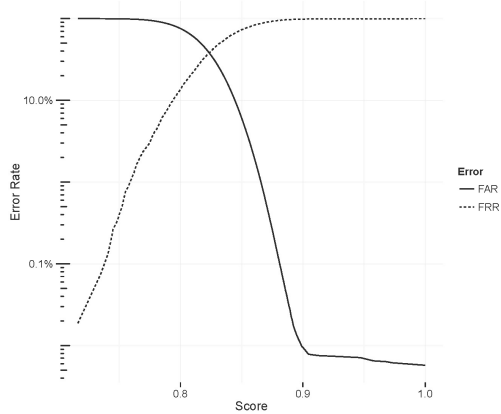
(b) S4F



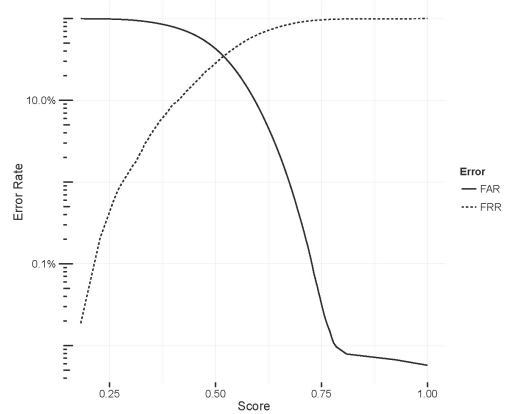
(c) CohortLDA



(d) LRPCA



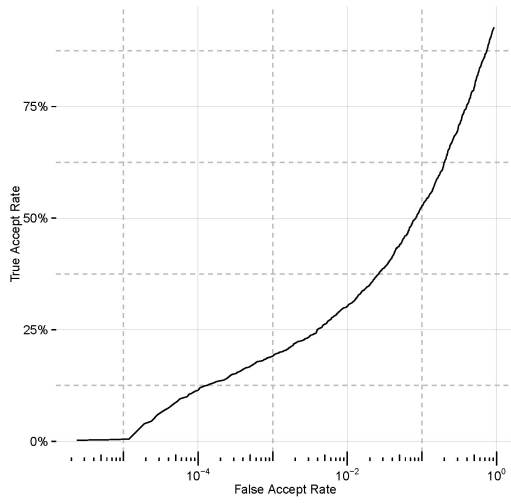
(e) LDA



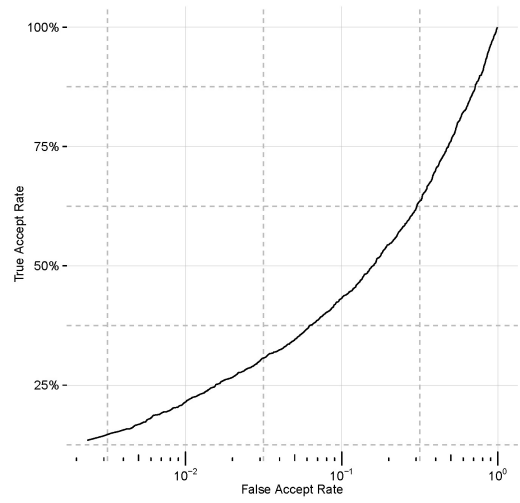
(f) PCA

APPENDICES B

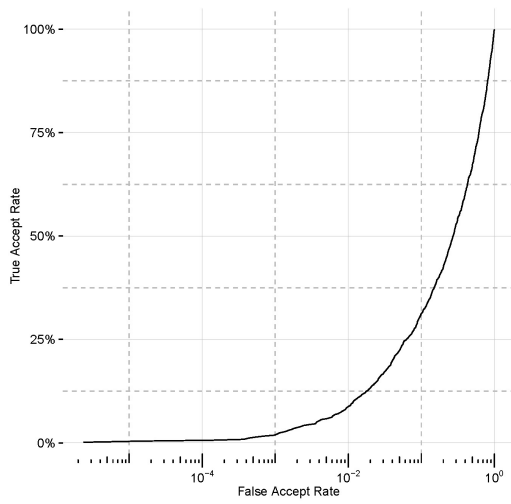
Young-to-Old Identification Results



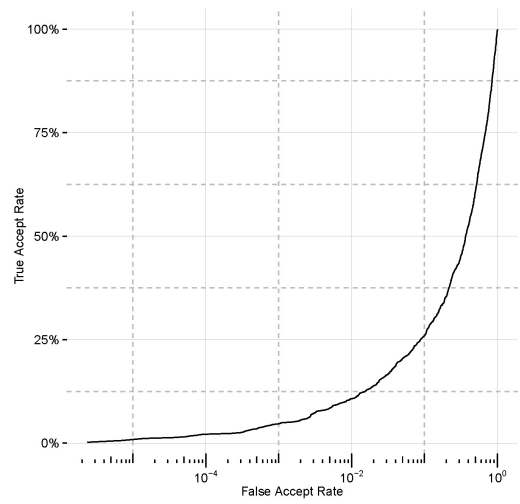
(a) Cognittec



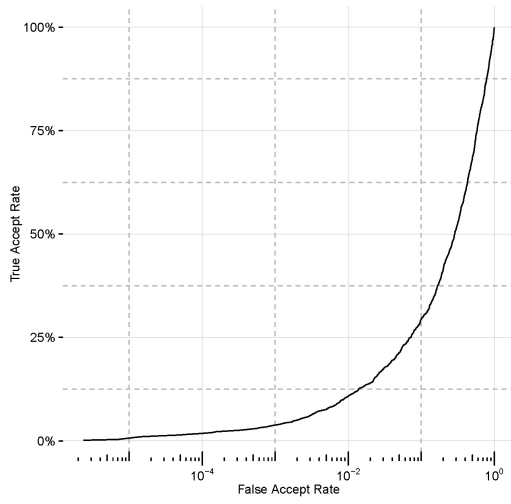
(b) S4F



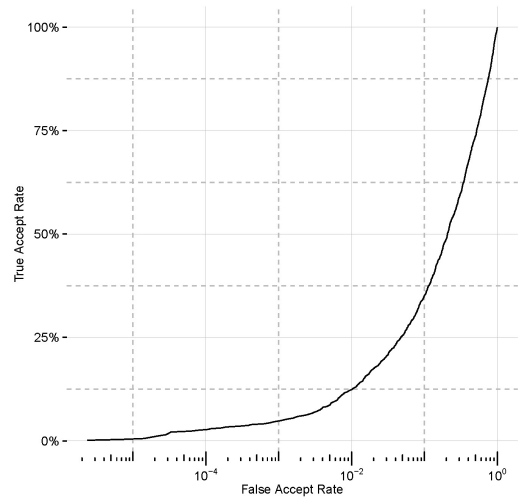
(c) CohortLDA



(d) LRPCA

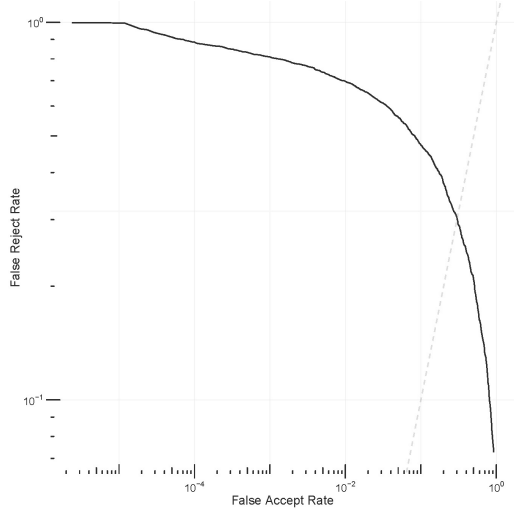


(e) LDA

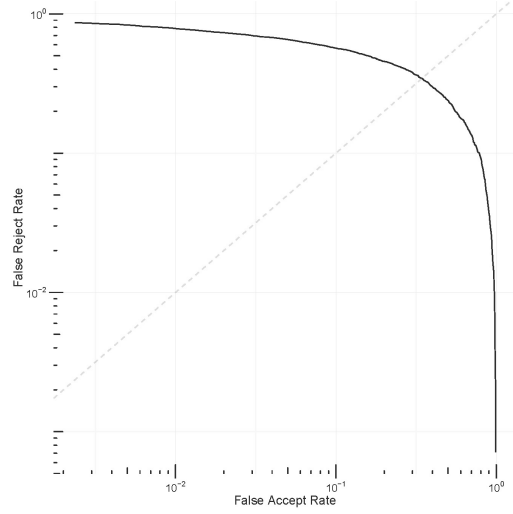


(f) PCA

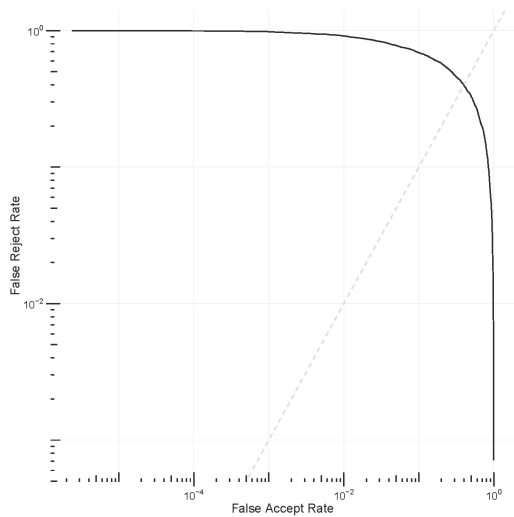
Figure 6.21: Photo-Tagging Scenario - ROC



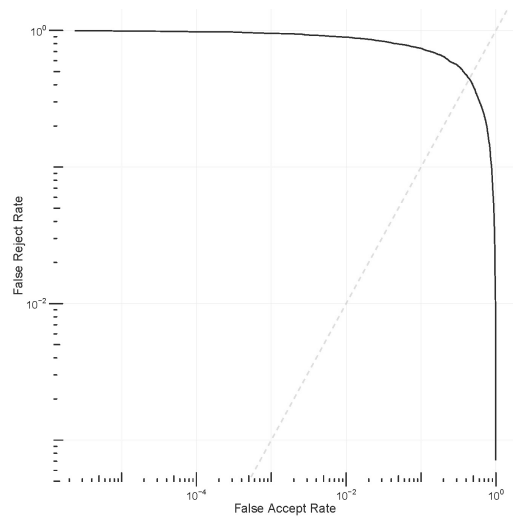
(a) Cognitec



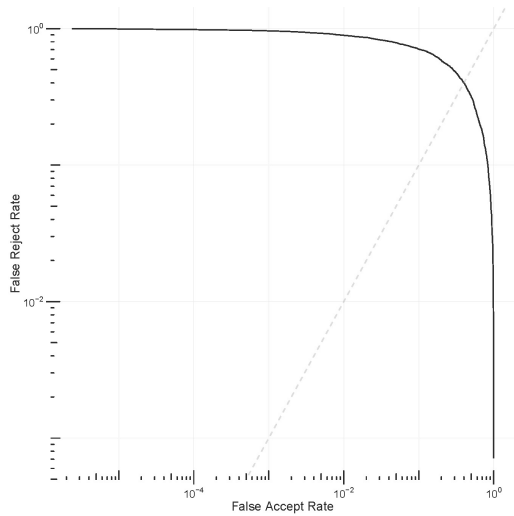
(b) S4F



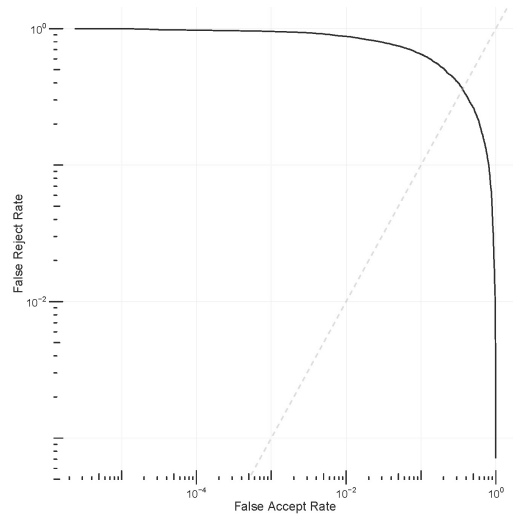
(c) CohortLDA



(d) LRPCA

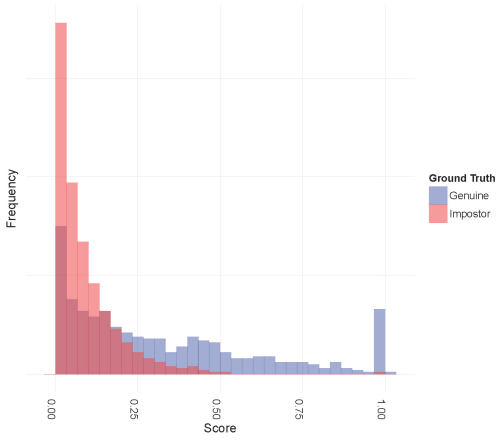


(e) LDA

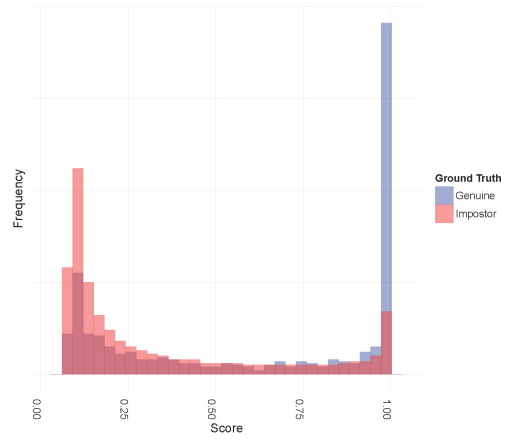


(f) PCA

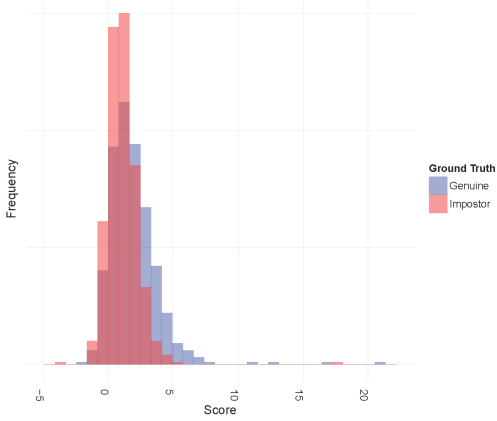
Figure 6.22: Photo-Tagging Scenario - DET



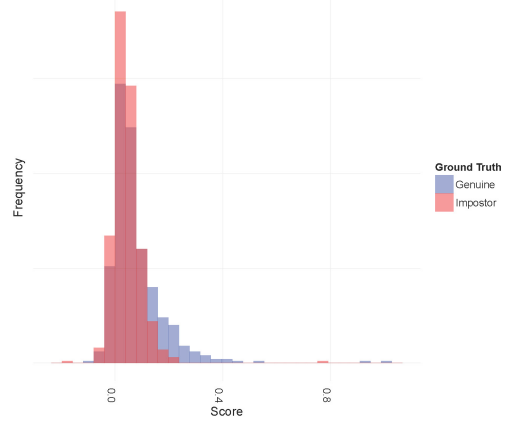
(a) Cognitec



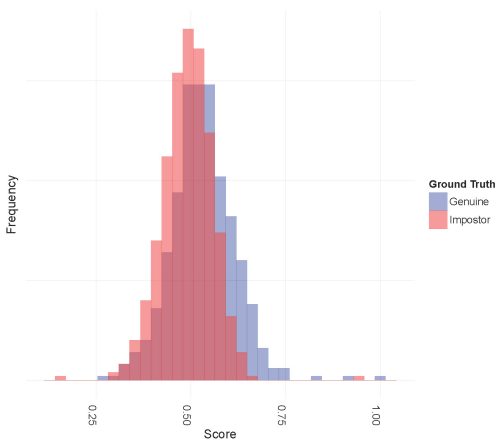
(b) S4F



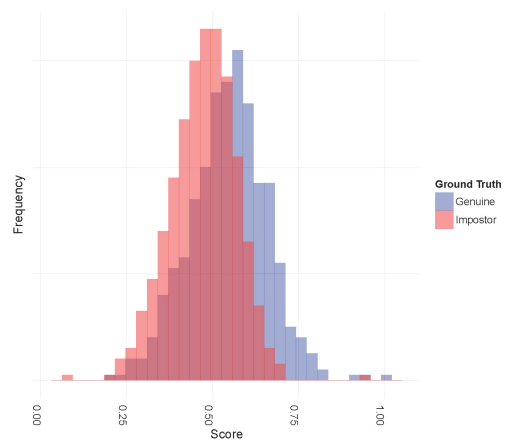
(c) CohortLDA



(d) LRPCA

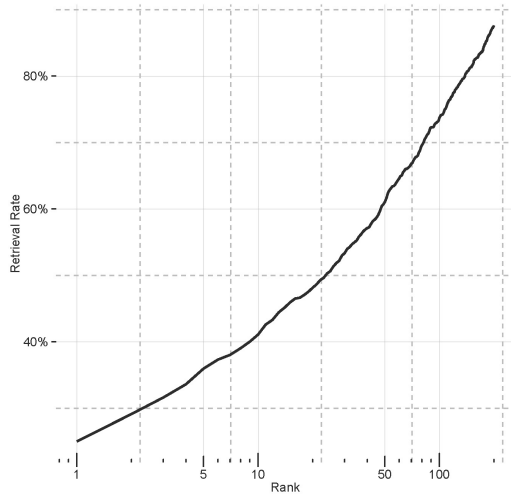


(e) LDA

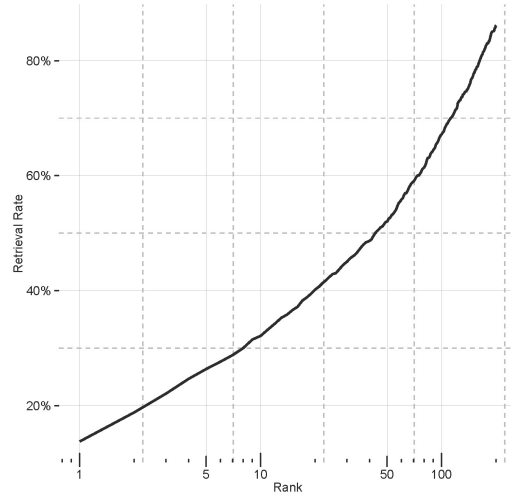


(f) PCA

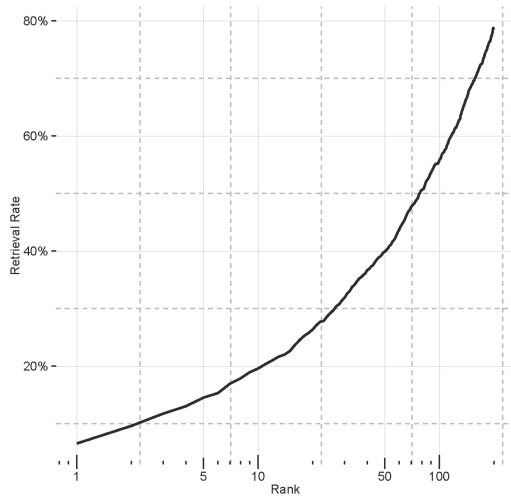
Figure 6.23: Photo-Tagging Scenario - Score Histogram



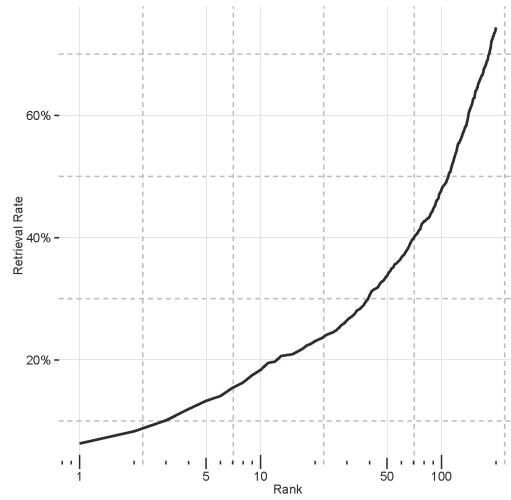
(a) Cognitec



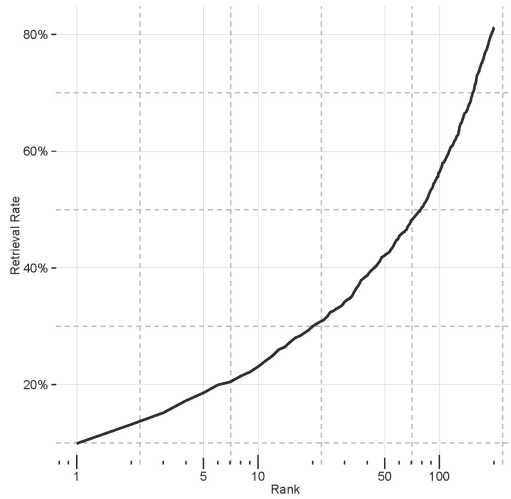
(b) S4F



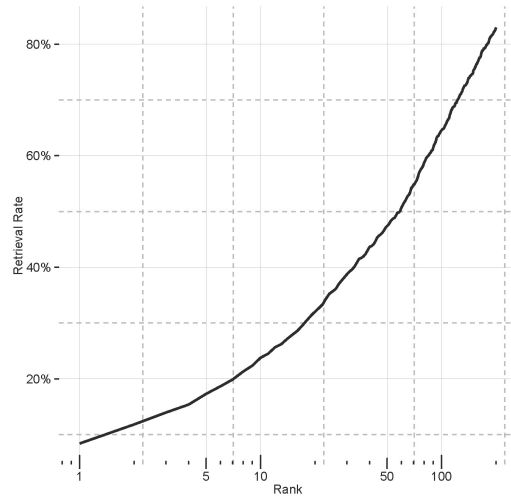
(c) CohortLDA



(d) LRPCA

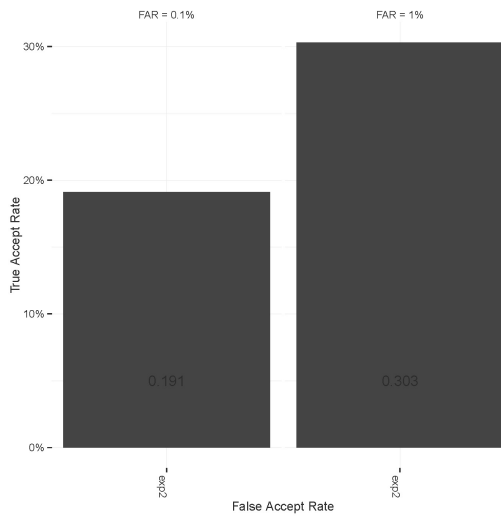


(e) LDA

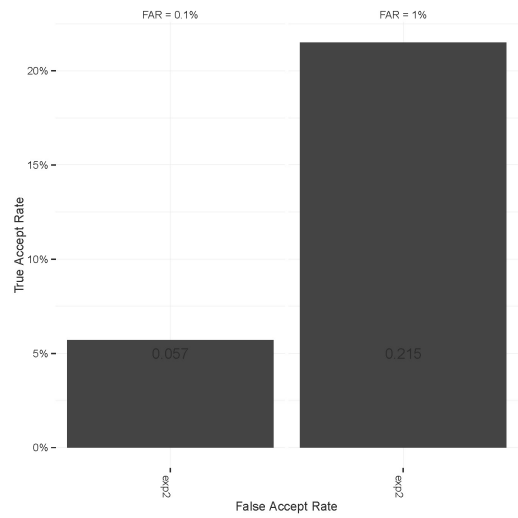


(f) PCA

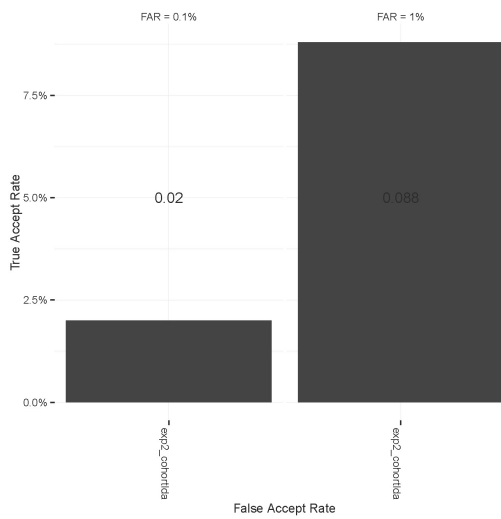
Figure 6.24: Photo-Tagging Scenario - CMC



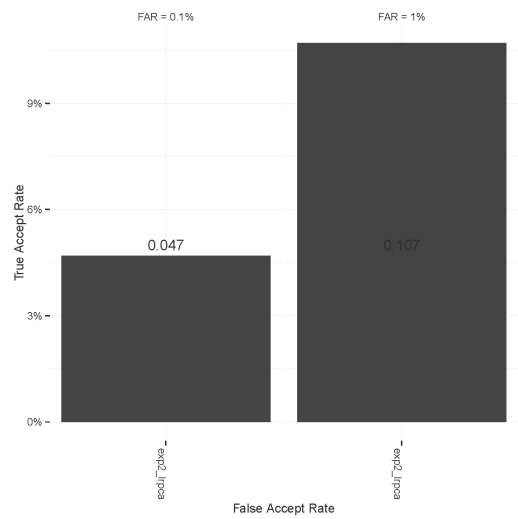
(a) Cognitec



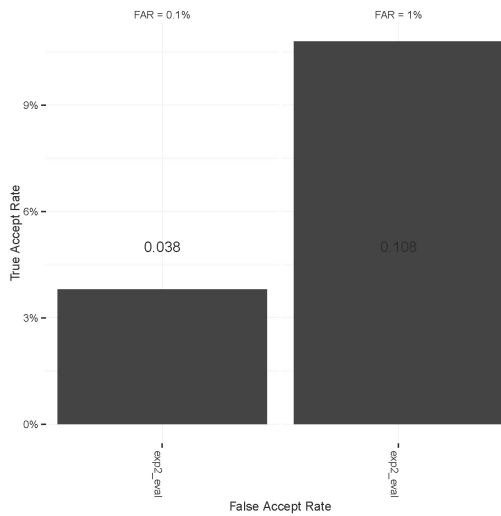
(b) S4F



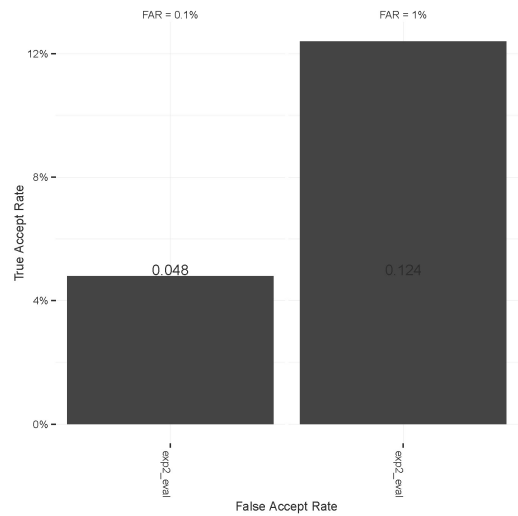
(c) CohortLDA



(d) LRPCA

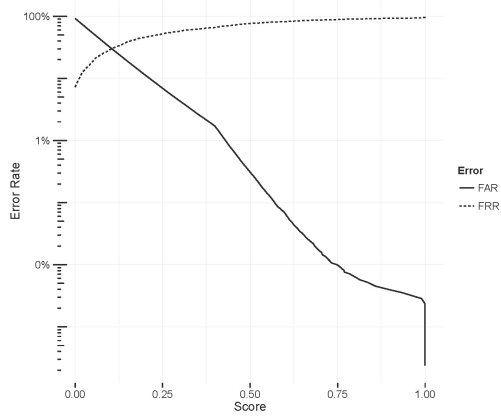


(e) LDA

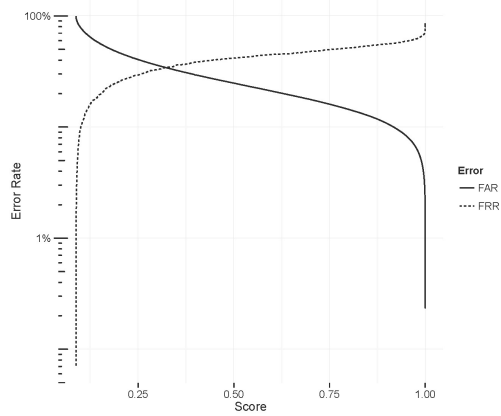


(f) PCA

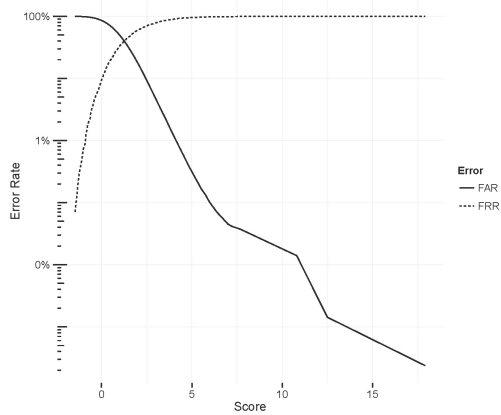
Figure 6.25: Photo-Tagging Scenario - ROC Scores



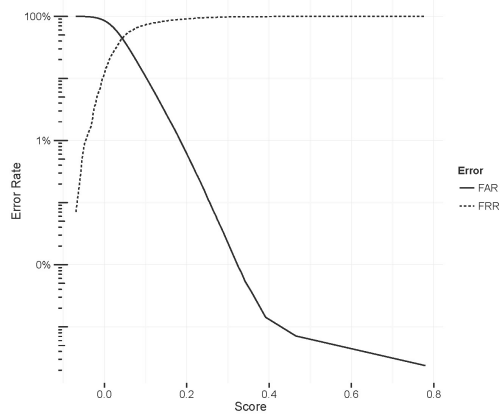
(a) Cognitec



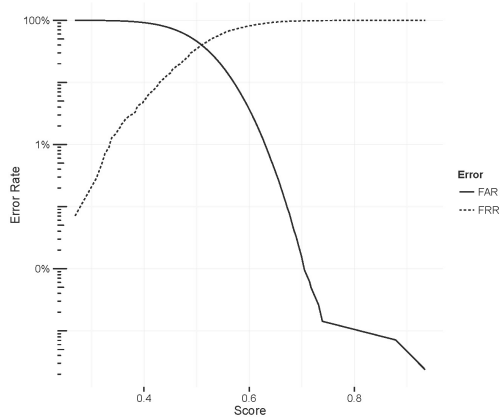
(b) S4F



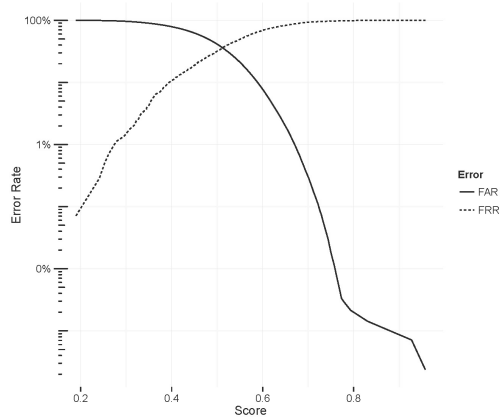
(c) CohortLDA



(d) LRPCA



(e) LDA

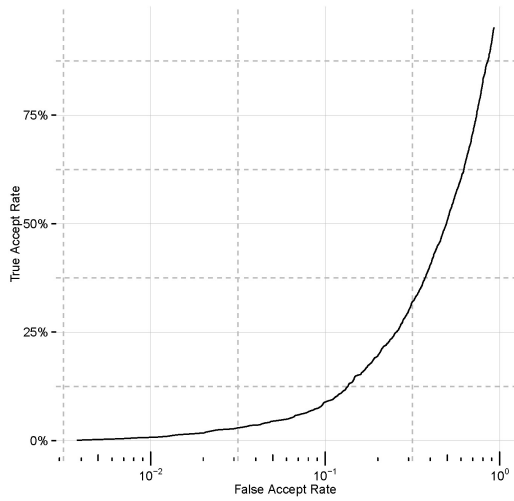


(f) PCA

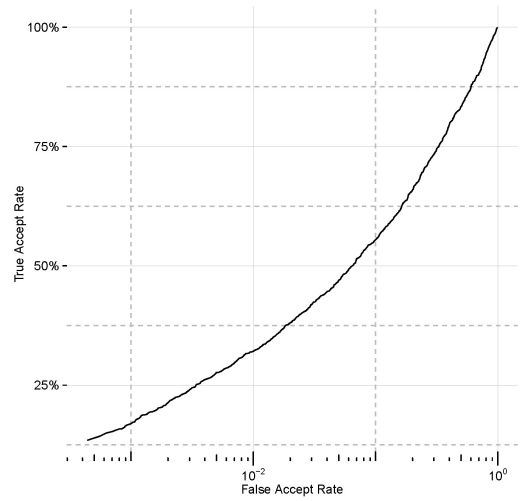
Figure 6.26: Photo-Tagging Scenario - EER

APPENDICES C

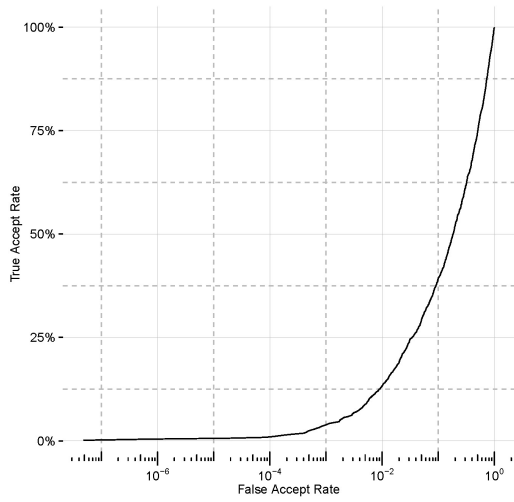
Augmented Young-to-Old Identification Results



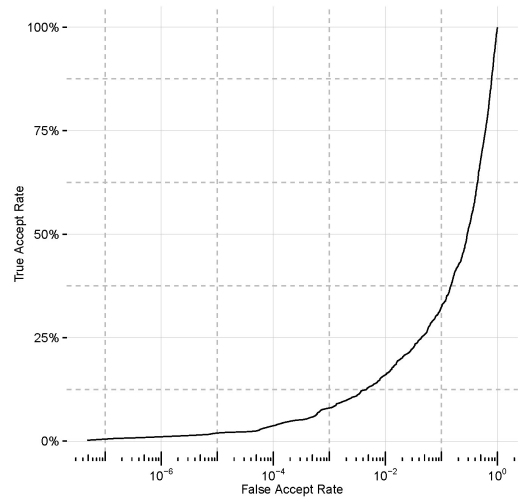
(a) Cognitec



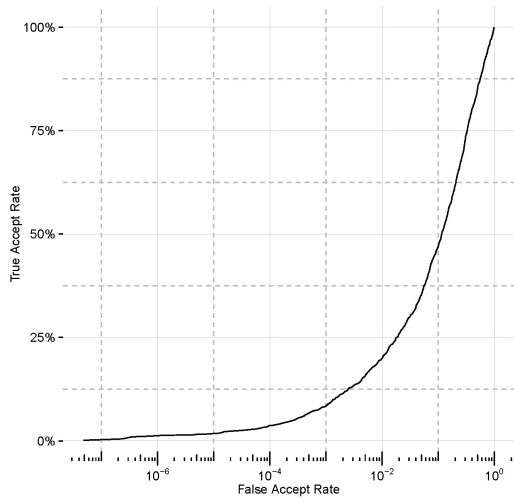
(b) S4F



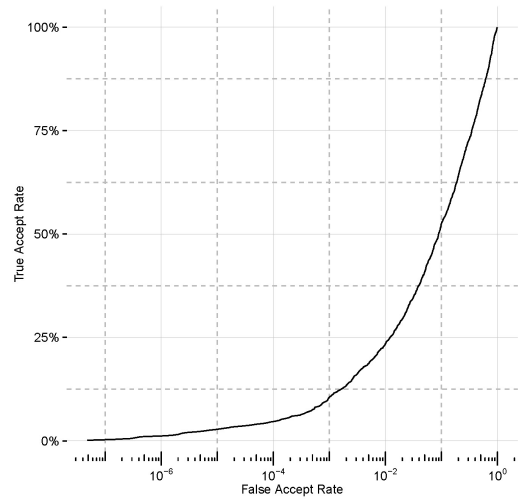
(c) CohortLDA



(d) LRPCA

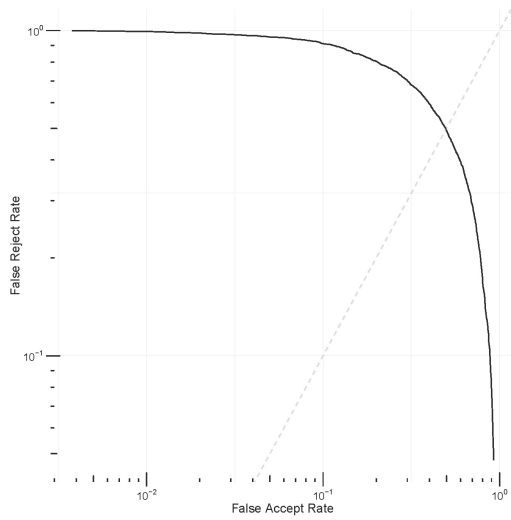


(e) LDA

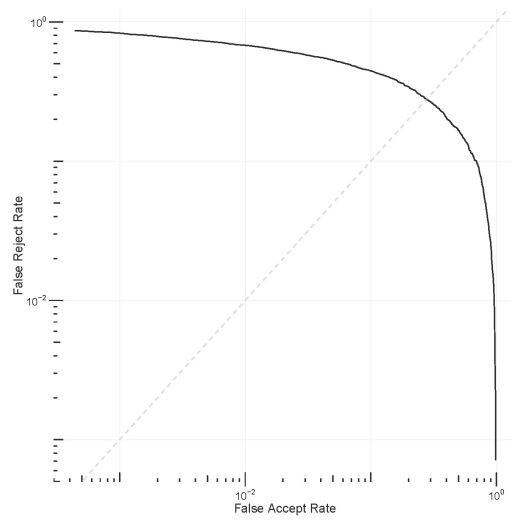


(f) PCA

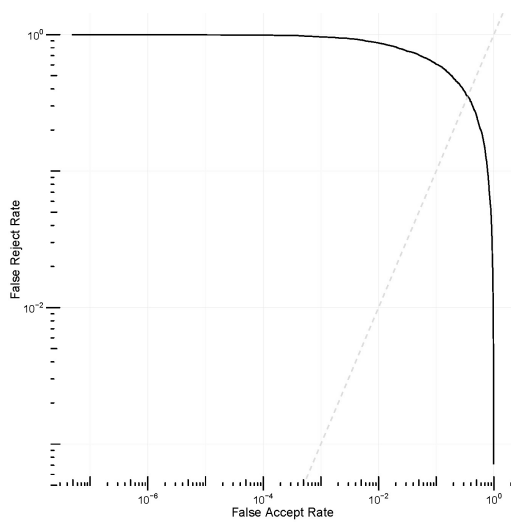
Figure 6.27: Augmented Photo-Tagging Scenario - ROC



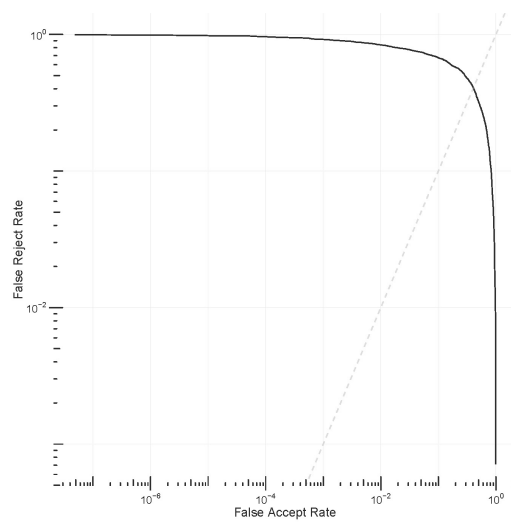
(a) Cognitec



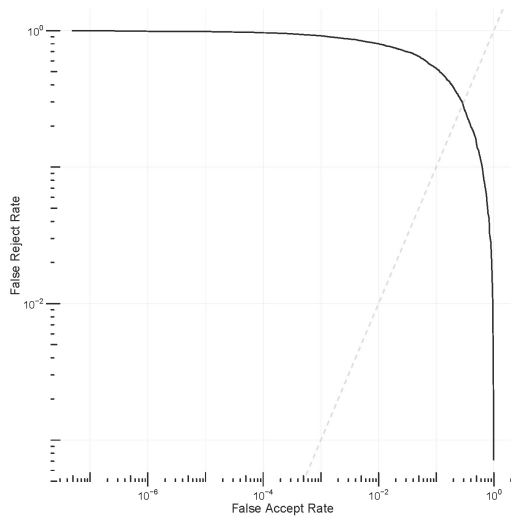
(b) S4F



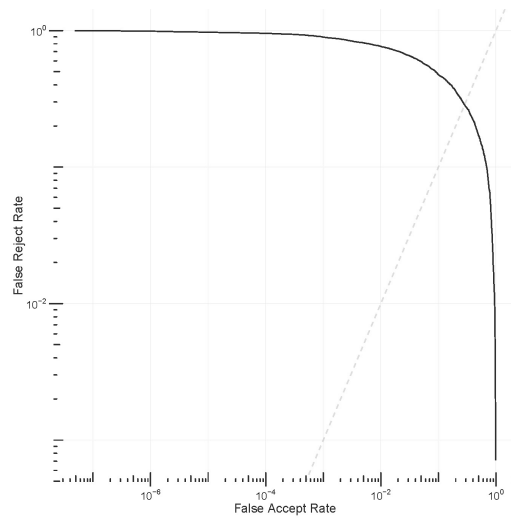
(c) CohortLDA



(d) LRPCA

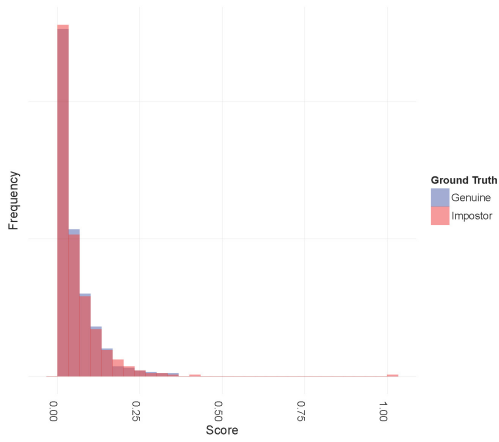


(e) LDA

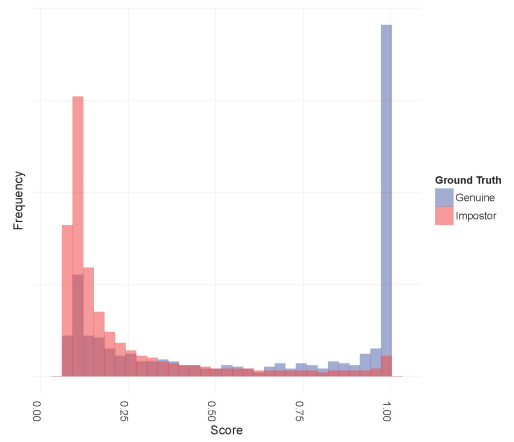


(f) PCA

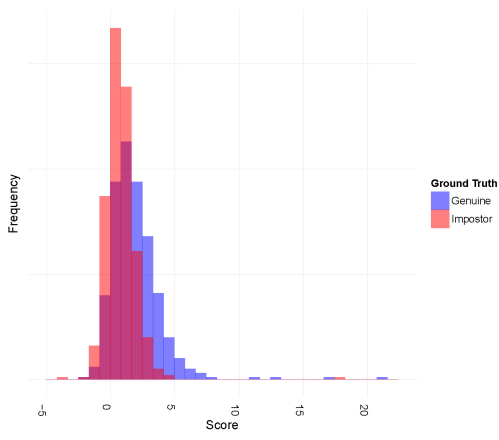
Figure 6.28: Augmented Photo-Tagging Scenario - DET



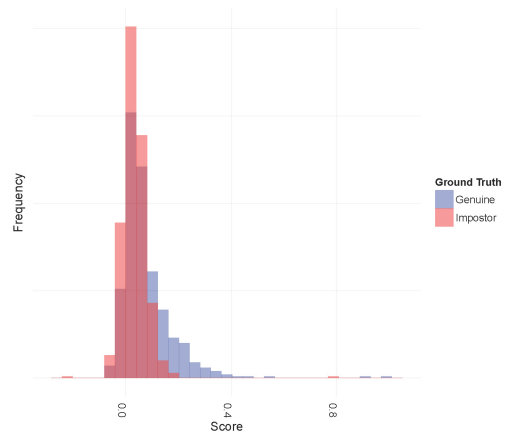
(a) Cognitec



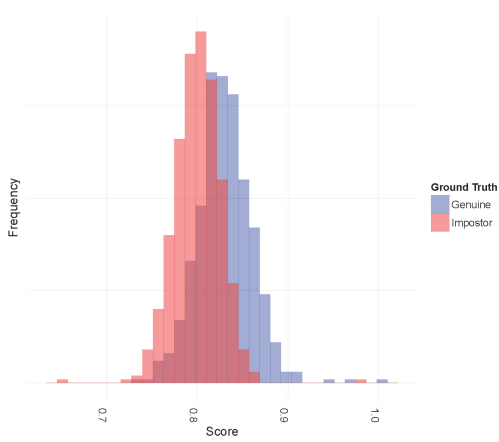
(b) S4F



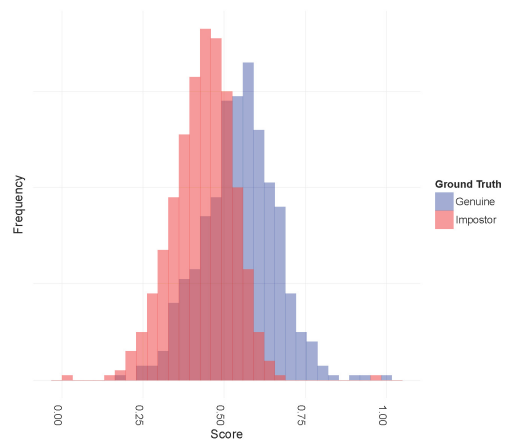
(c) CohortLDA



(d) LRPCA

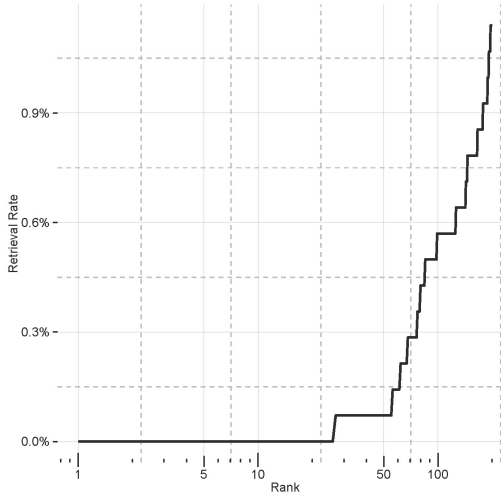


(e) LDA

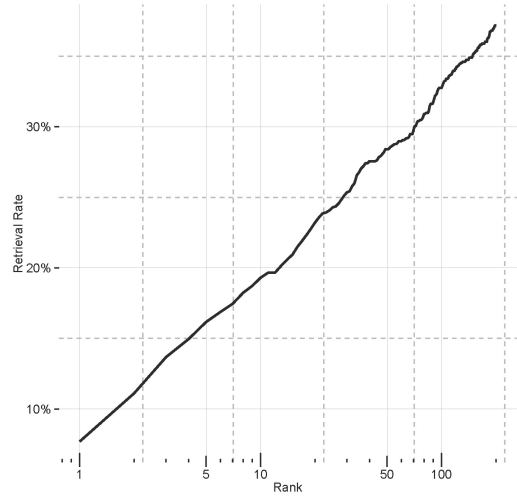


(f) PCA

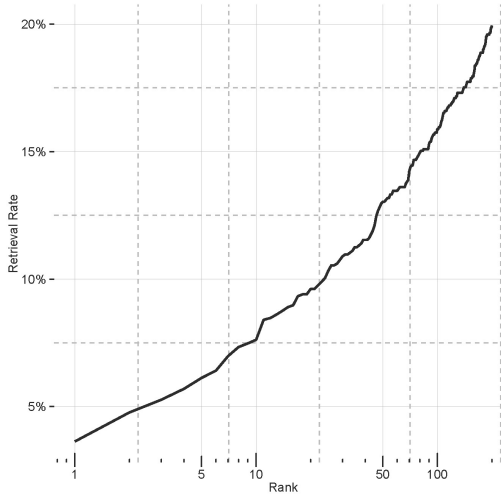
Figure 6.29: Augmented Photo-Tagging Scenario - Score Histogram



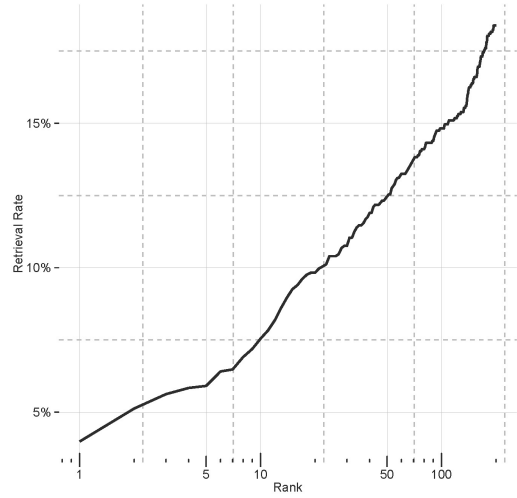
(a) Cognitec



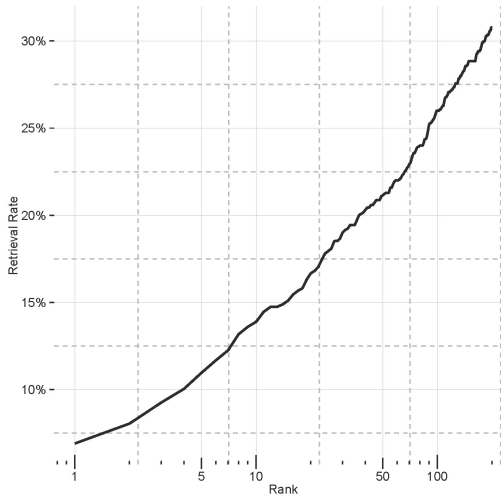
(b) S4F



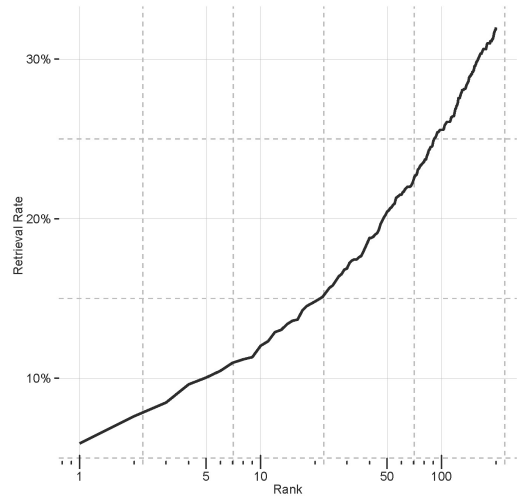
(c) CohortLDA



(d) LRPCA

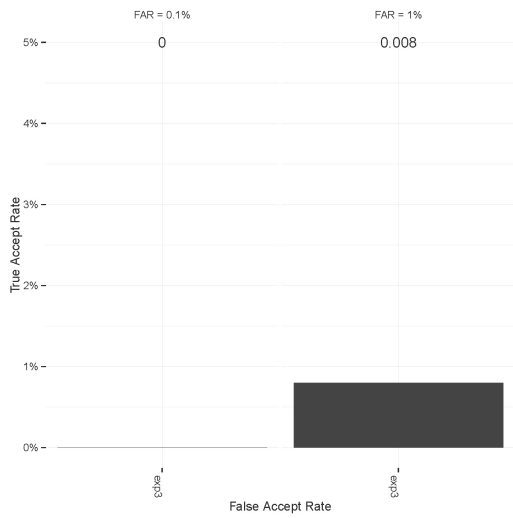


(e) LDA

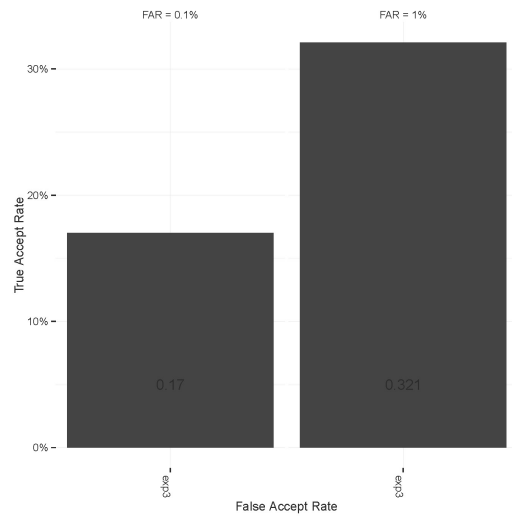


(f) PCA

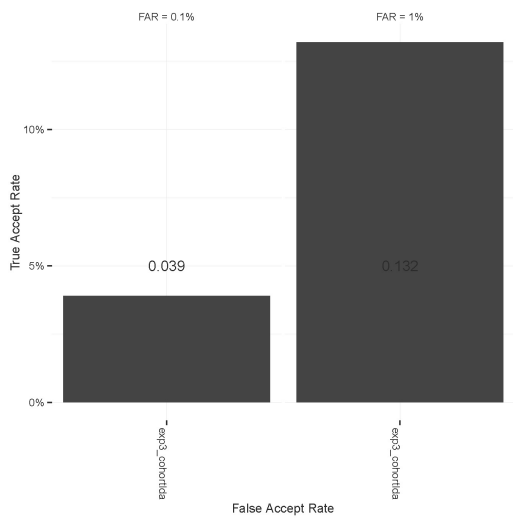
Figure 6.30: Augmented Photo-Tagging Scenario - CMC



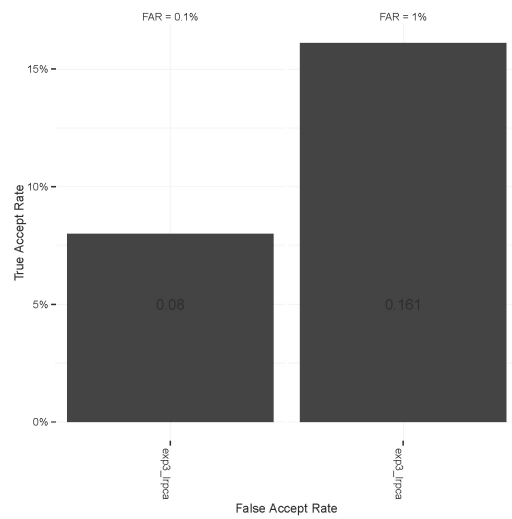
(a) Cognitex



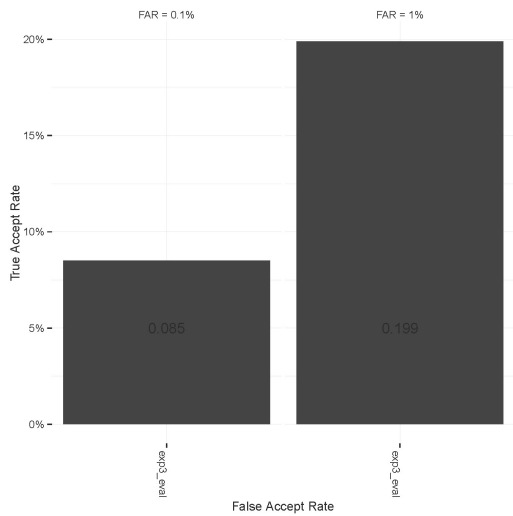
(b) S4F



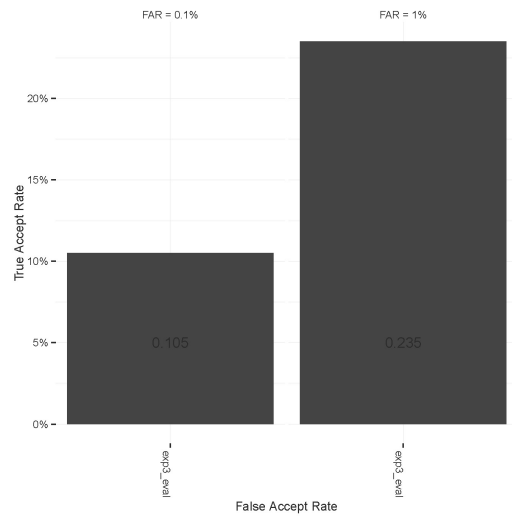
(c) CohortLDA



(d) LRPCA

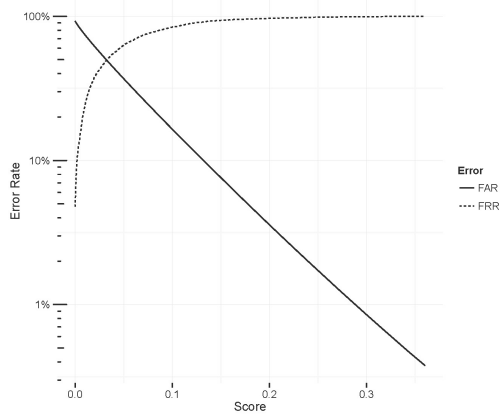


(e) LDA

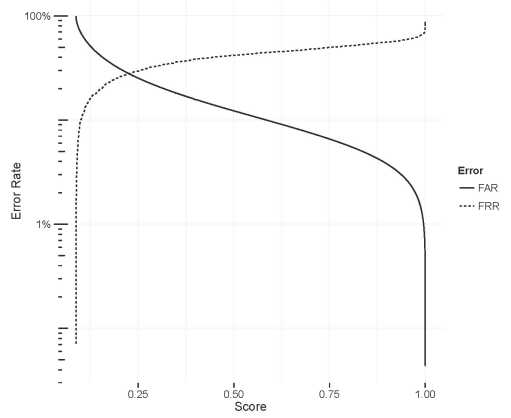


(f) PCA

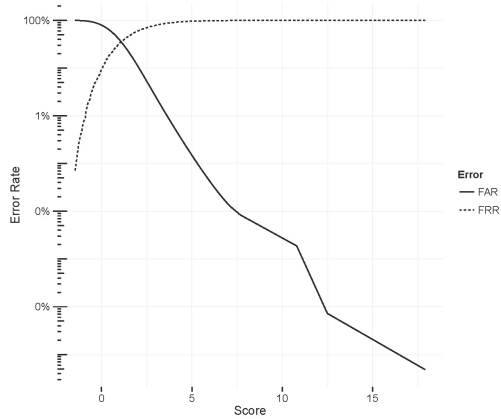
Figure 6.31: Augmented Photo-Tagging Scenario - ROC Scores



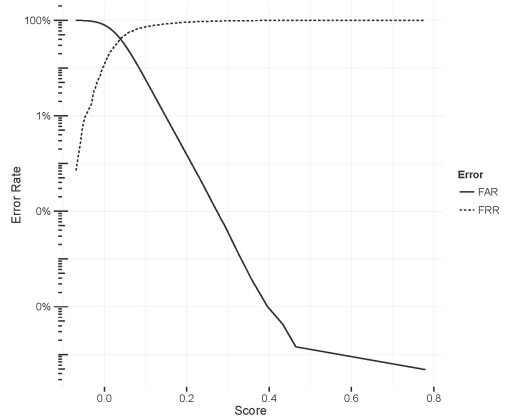
(a) Cognitec



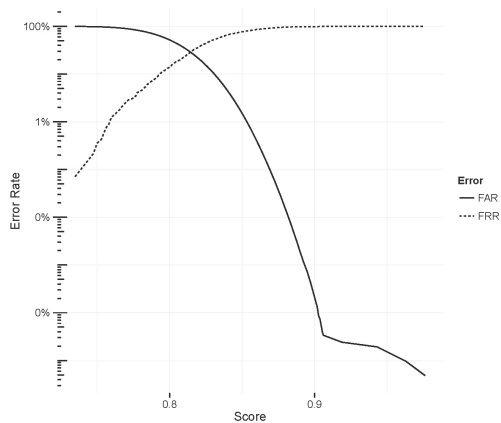
(b) S4F



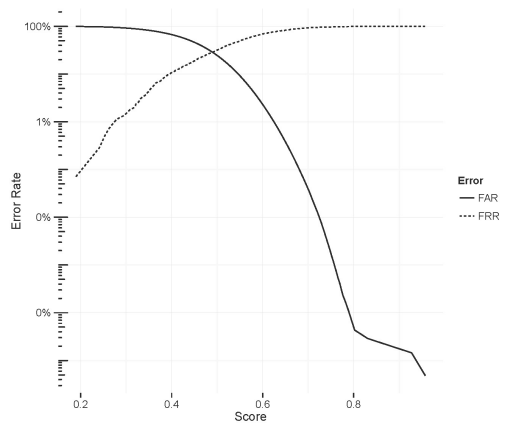
(c) CohortLDA



(d) LRPCA



(e) LDA

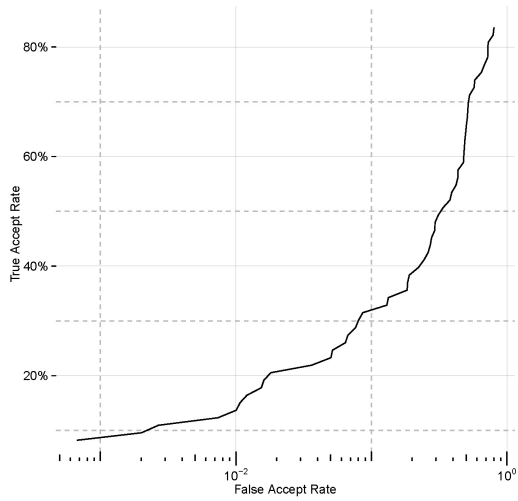


(f) PCA

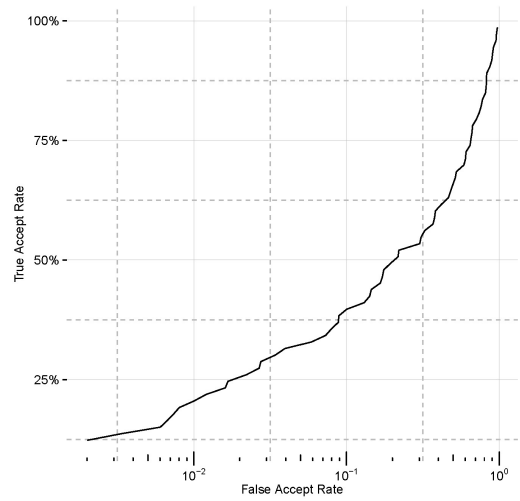
Figure 6.32: Augmented Photo-Tagging Scenario - EER

APPENDICES D

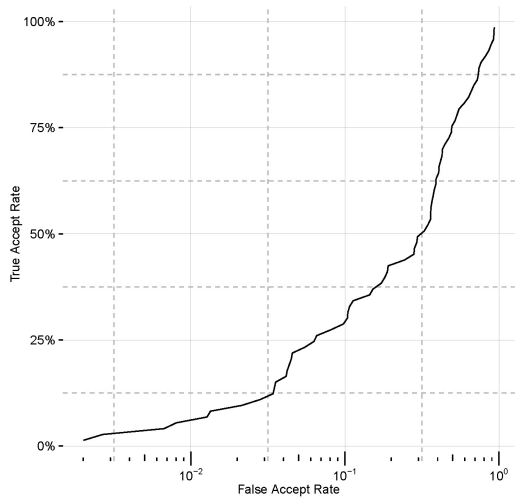
Age Gap 4 - 3-5 to 9-11 Results



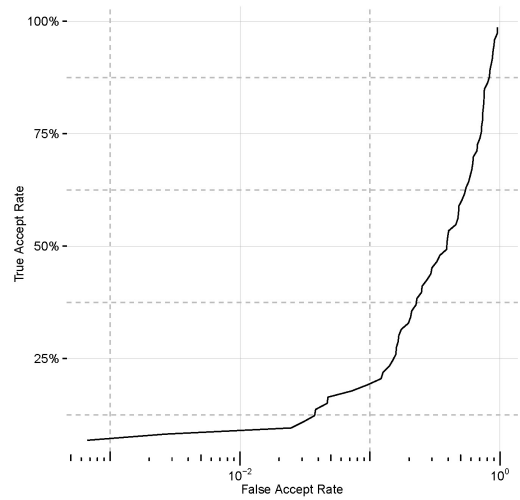
(a) Cognitec



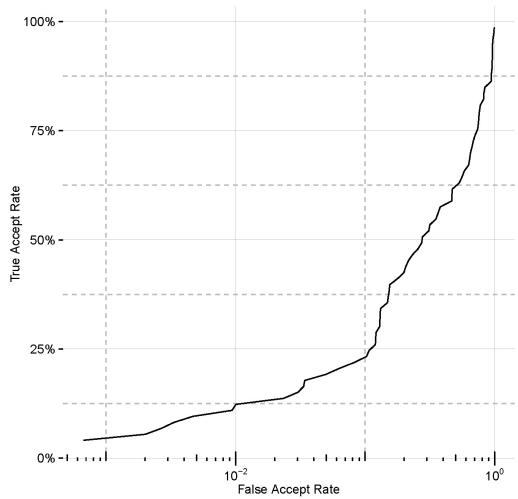
(b) S4F



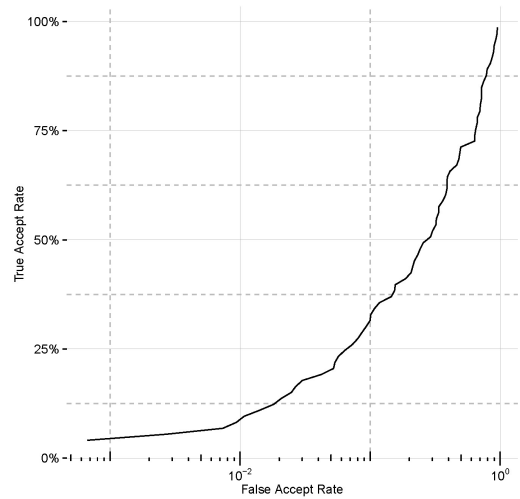
(c) CohortLDA



(d) LRPCA

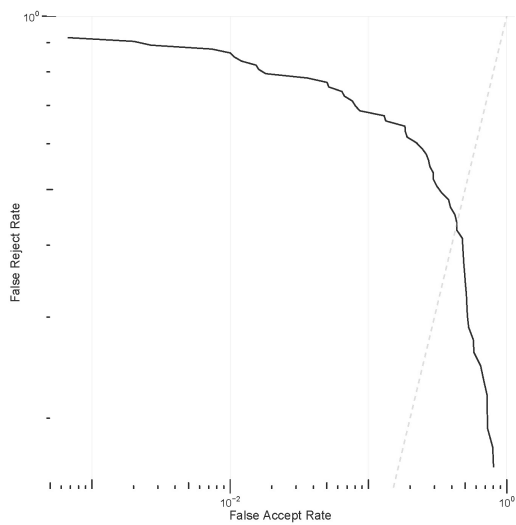


(e) LDA

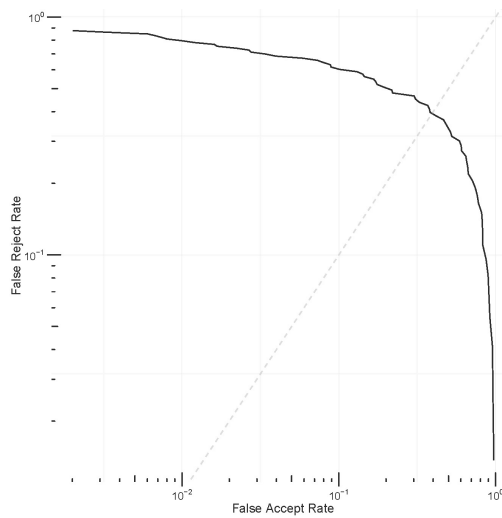


(f) PCA

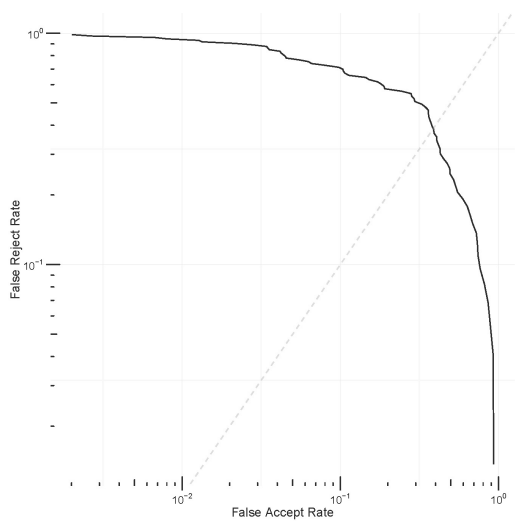
Figure 6.33: Age Gap 3-5 to 9-11 - ROC



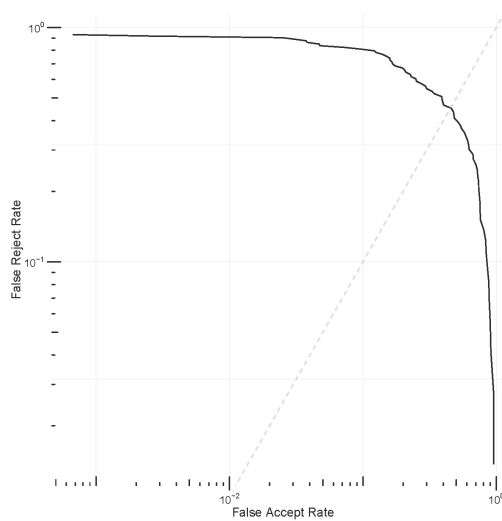
(a) Cognitec



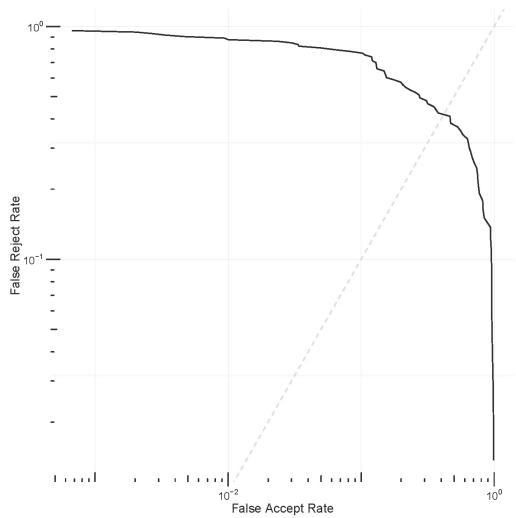
(b) S4F



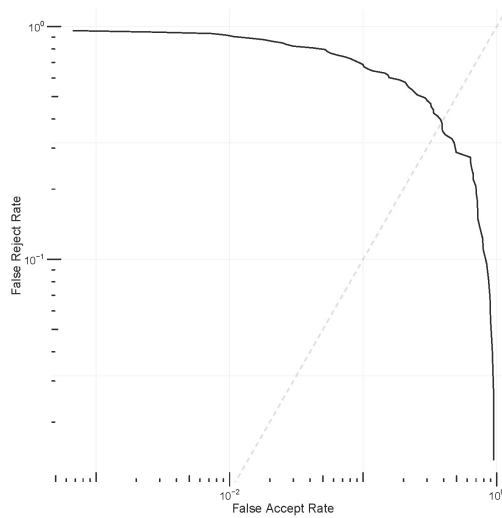
(c) CohortLDA



(d) LRPCA

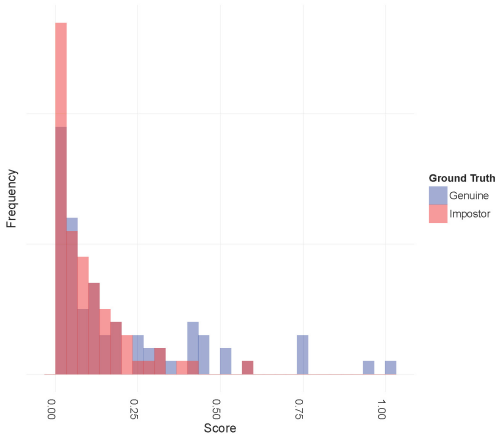


(e) LDA

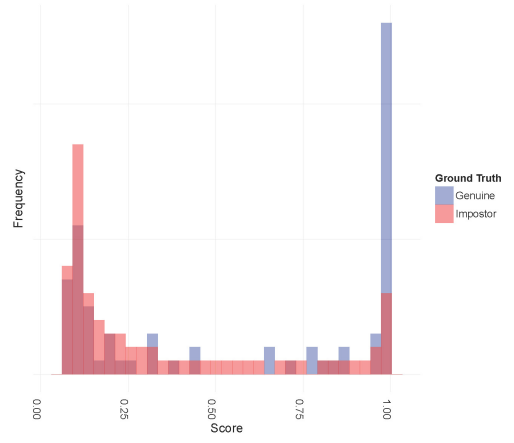


(f) PCA

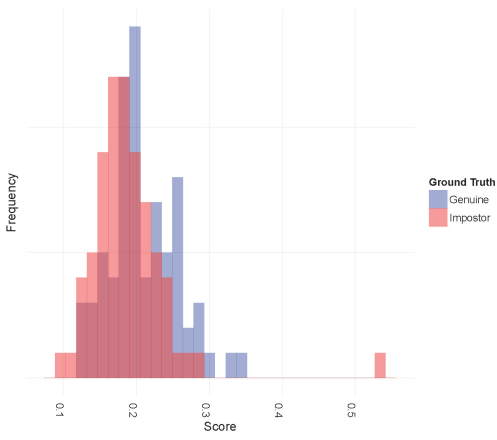
Figure 6.34: Age Gap 3-5 to 9-11 - DET



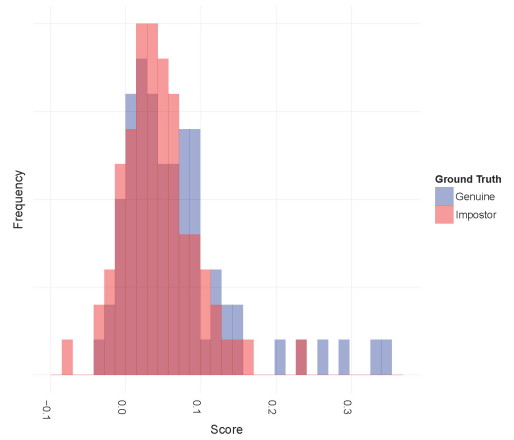
(a) Cognitec



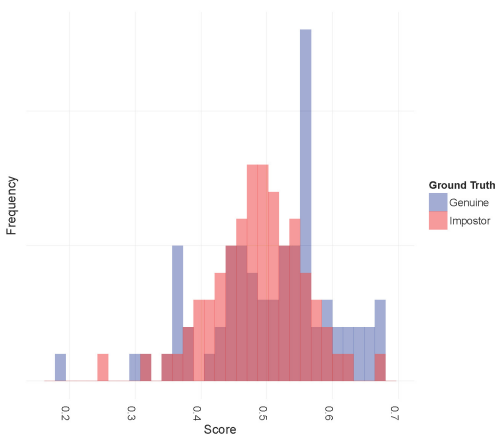
(b) S4F



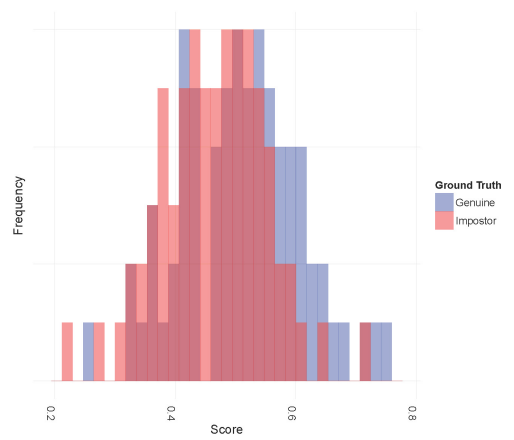
(c) CohortLDA



(d) LRPCA

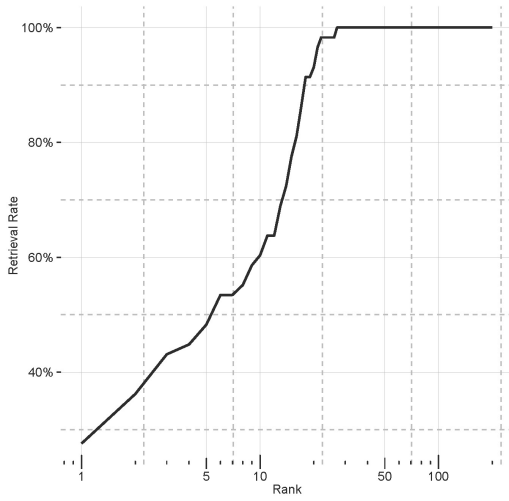


(e) LDA

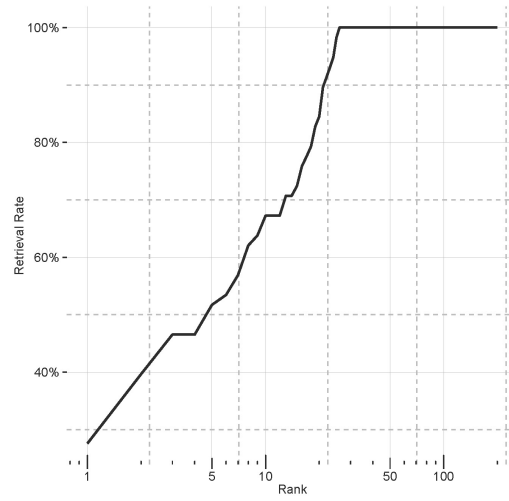


(f) PCA

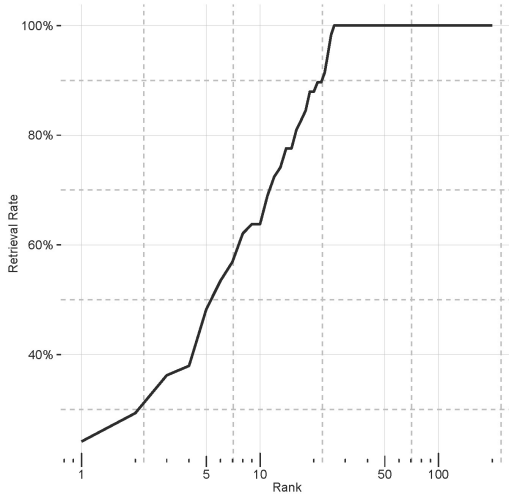
Figure 6.35: Age Gap 3-5 to 9-11 - Score Histogram



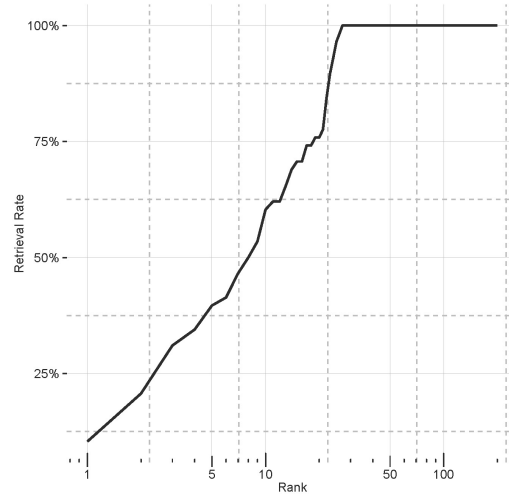
(a) Cognittec



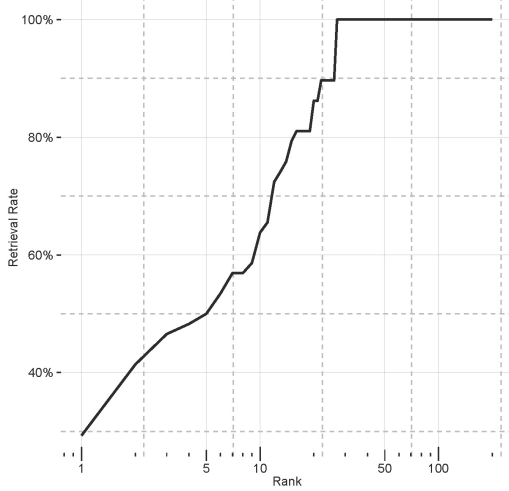
(b) S4F



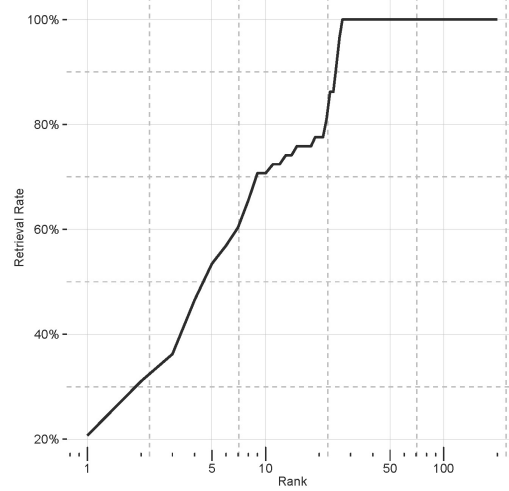
(c) CohortLDA



(d) LRPCA

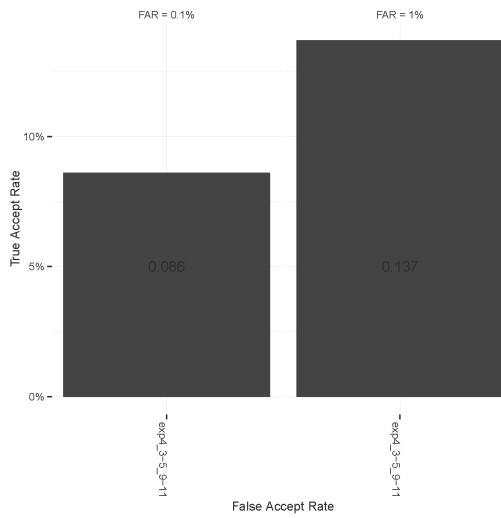


(e) LDA

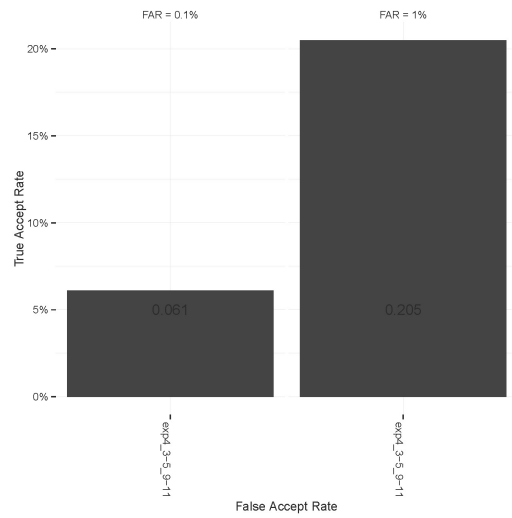


(f) PCA

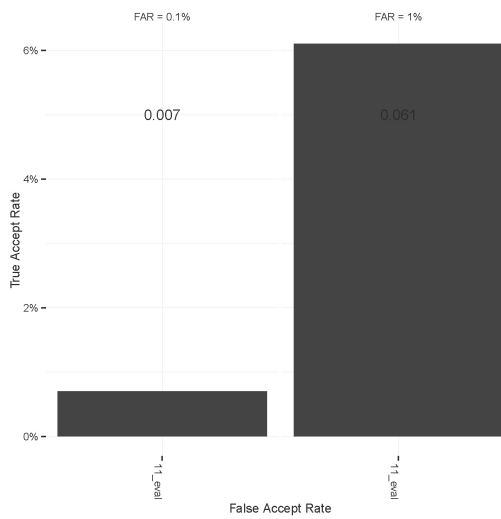
Figure 6.36: Age Gap 3-5 to 9-11 - CMC



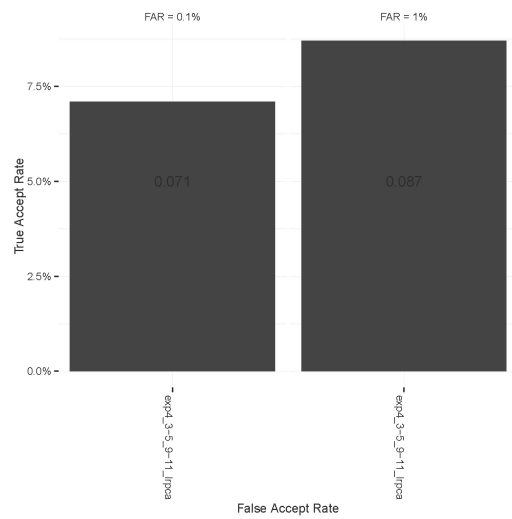
(a) Cognitec



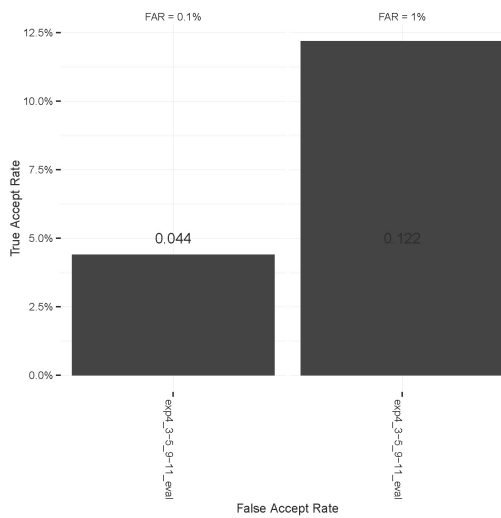
(b) S4F



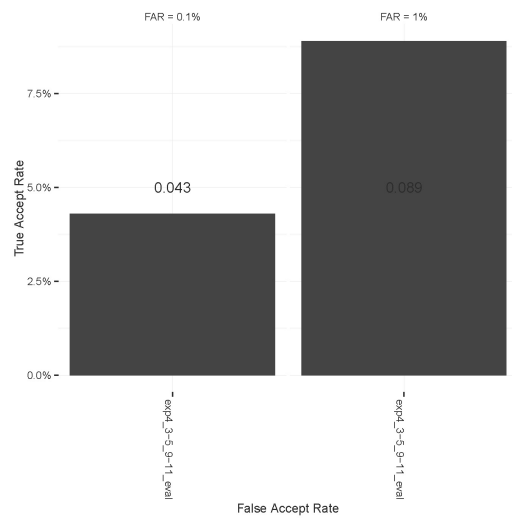
(c) CohortLDA



(d) LRPCA

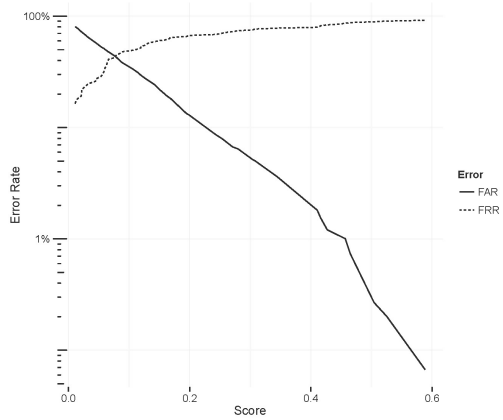


(e) LDA

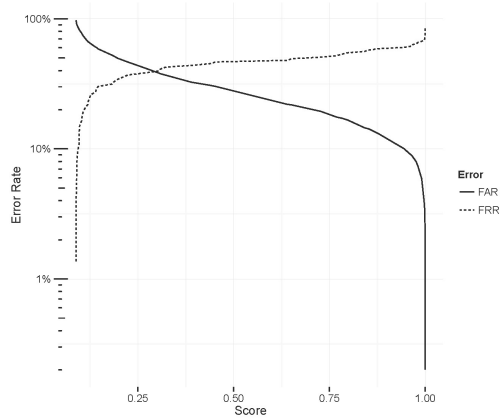


(f) PCA

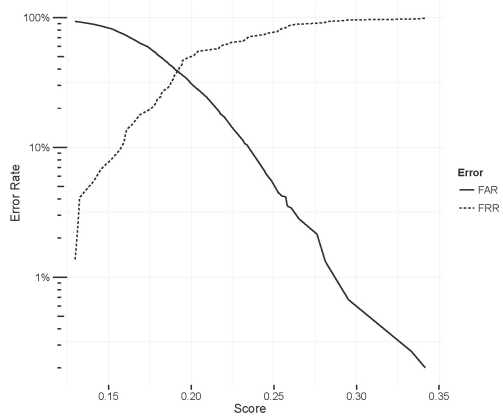
Figure 6.37: Age Gap 3-5 to 9-11 - ROC Scores



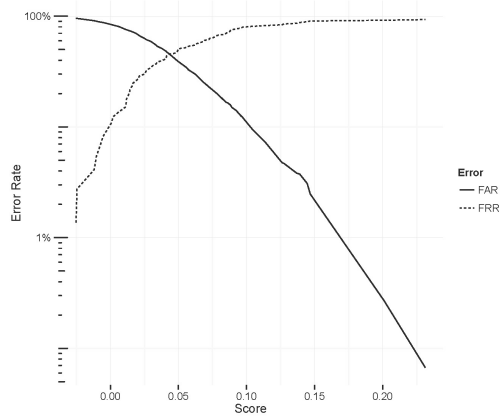
(a) Cognitec



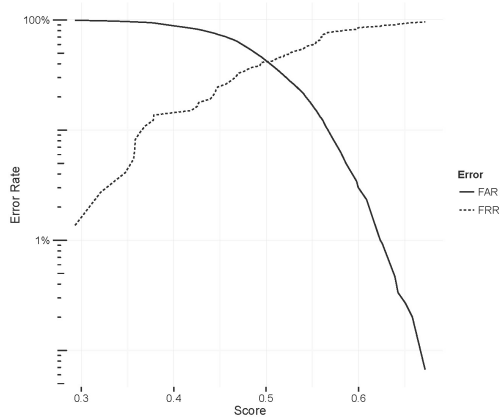
(b) S4F



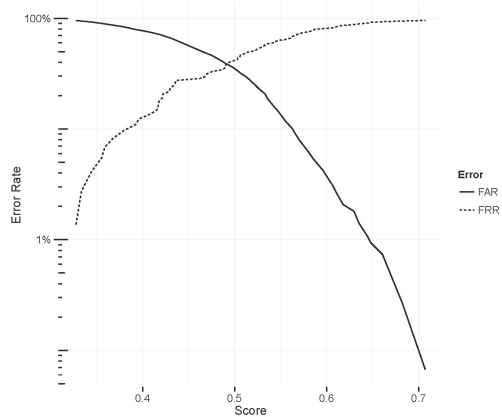
(c) CohortLDA



(d) LRPCA



(e) LDA

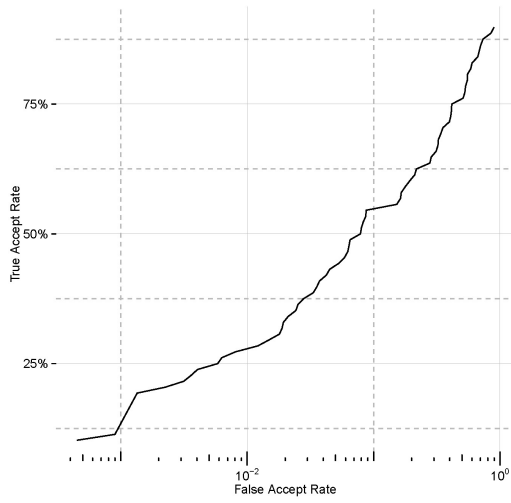


(f) PCA

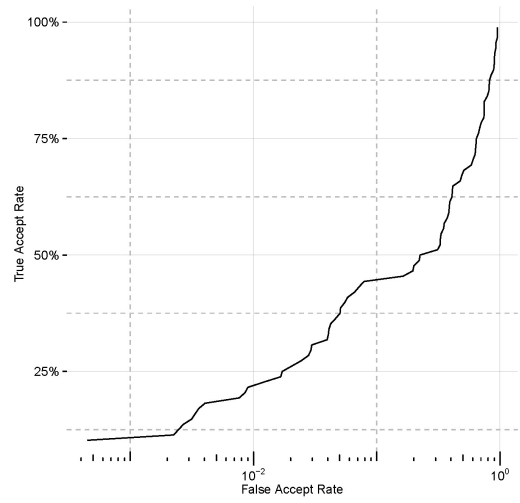
Figure 6.38: Age Gap 3-5 to 9-11 - EER

APPENDICES E

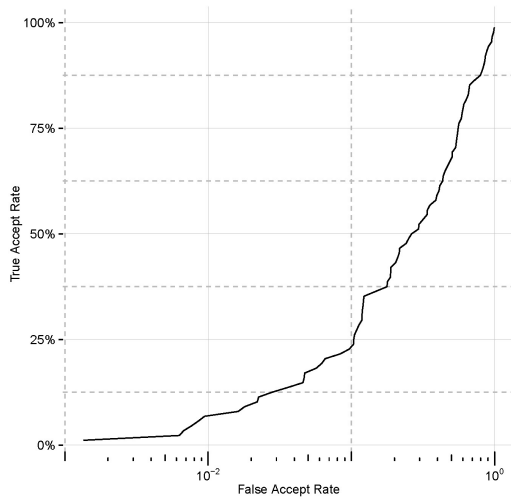
Experiment 4 - 3-5 to 12-14 Results



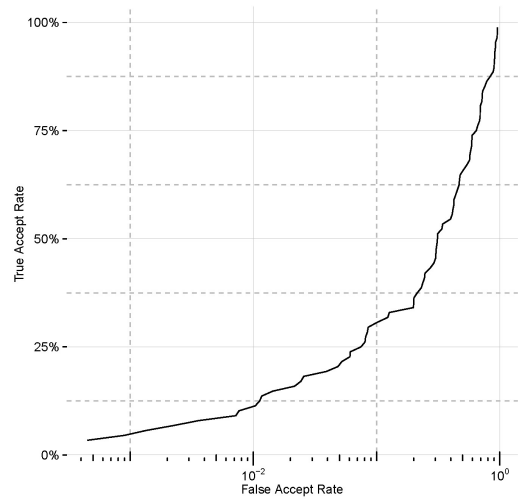
(a) Cognittec



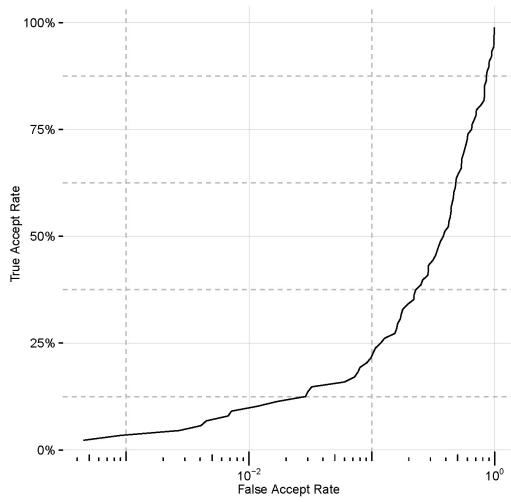
(b) S4F



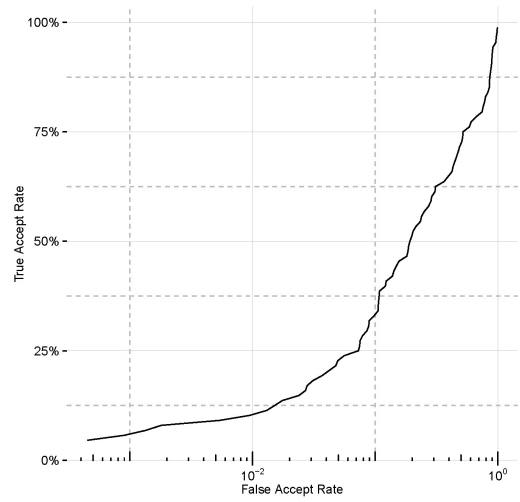
(c) cohortLDA



(d) LRPCA

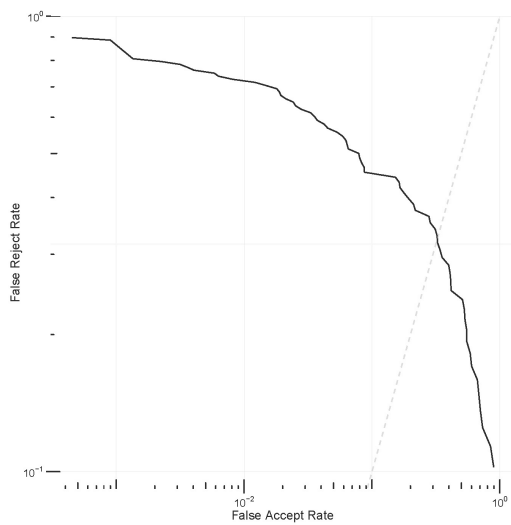


(e) LDA

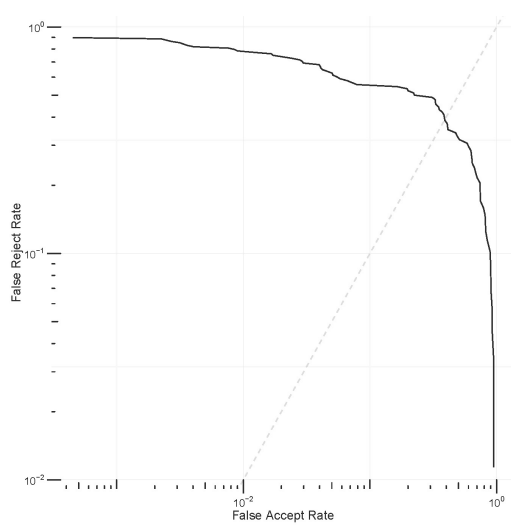


(f) PCA

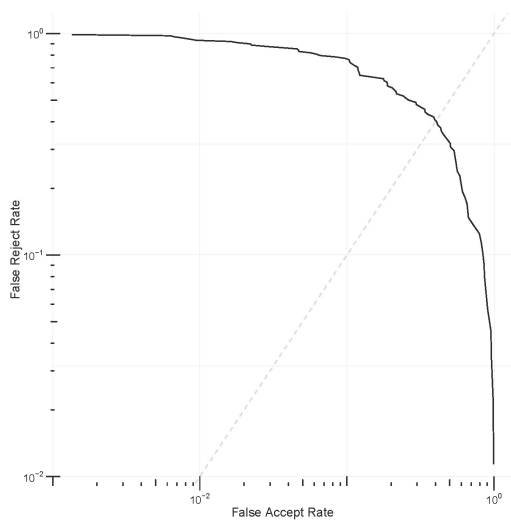
Figure 6.39: Experiment 4 3-5 to 12-14 - ROC



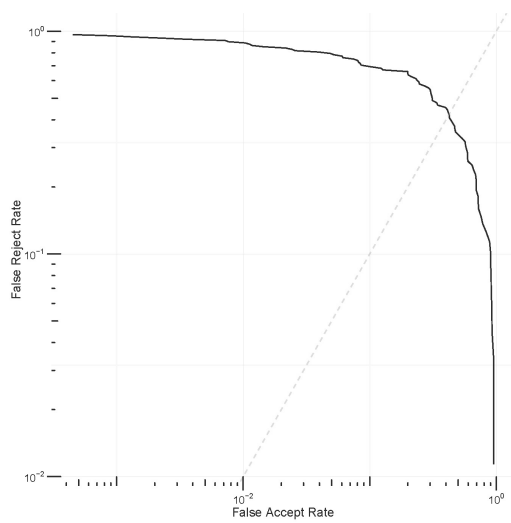
(a) Cognitec



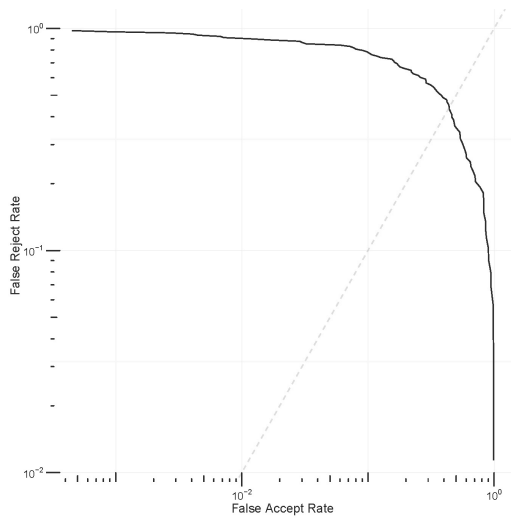
(b) S4F



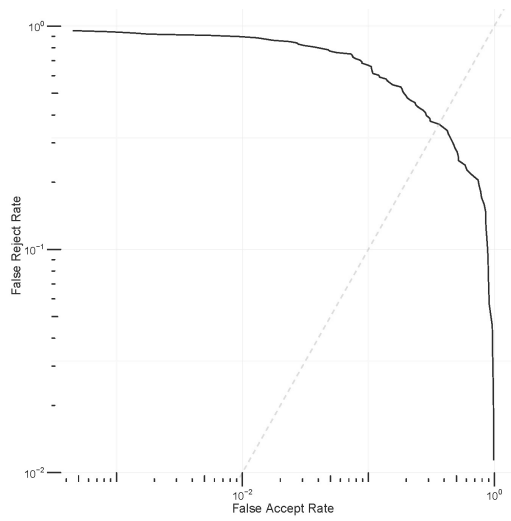
(c) cohortLDA



(d) LRPCA

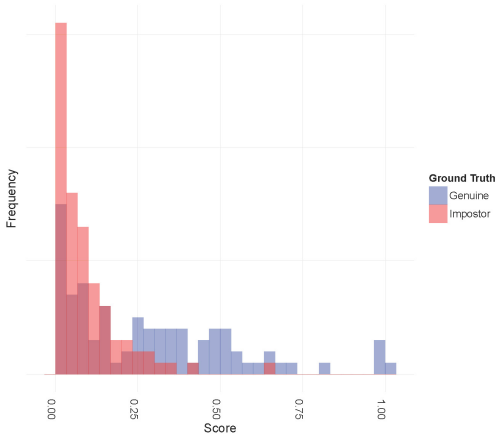


(e) LDA

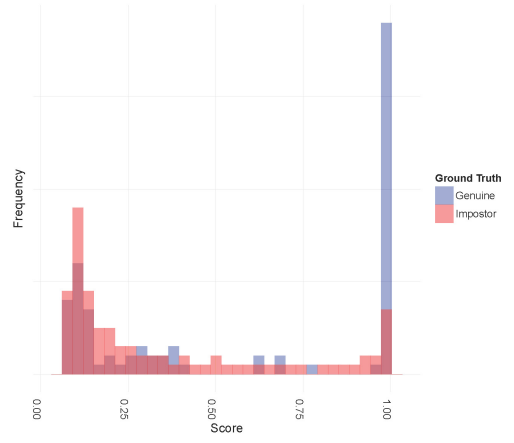


(f) PCA

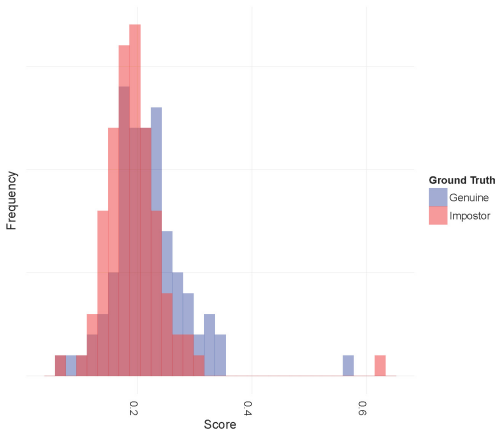
Figure 6.40: Experiment 4 3-5 to 12-14 - DET



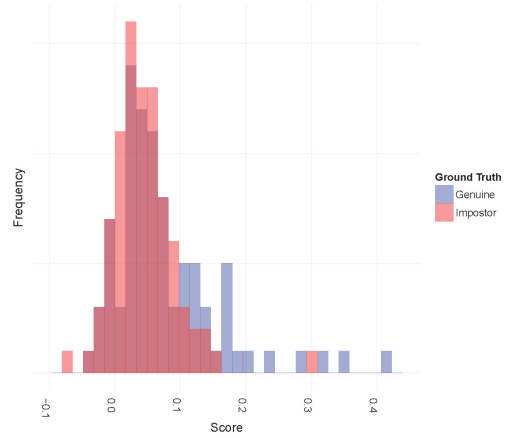
(a) Cognitec



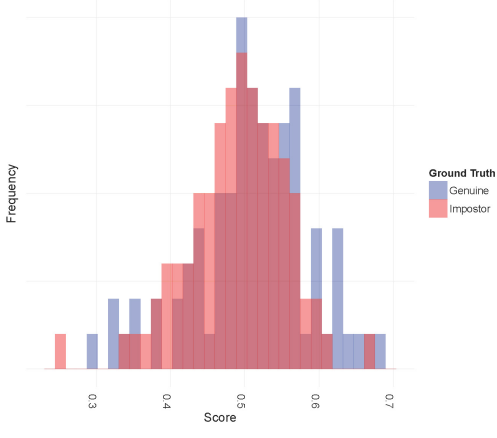
(b) S4F



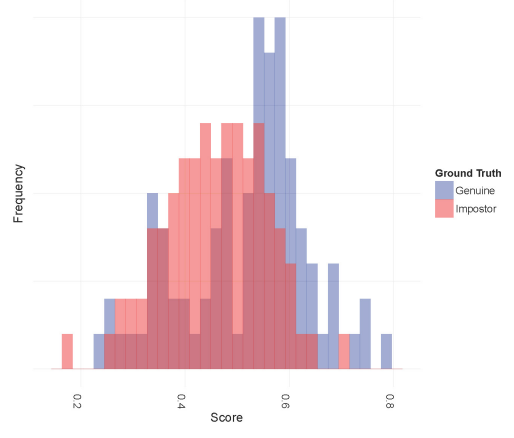
(c) cohortLDA



(d) LRPCA

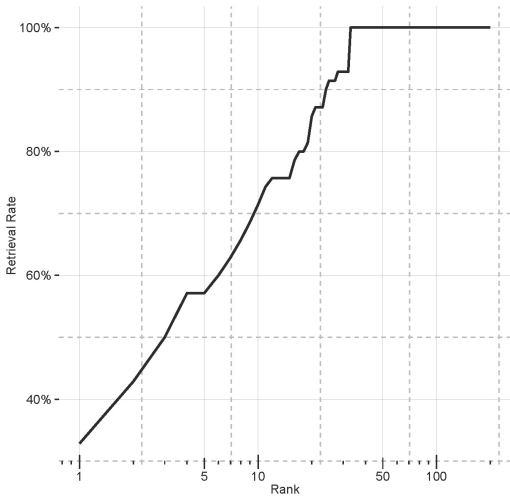


(e) LDA

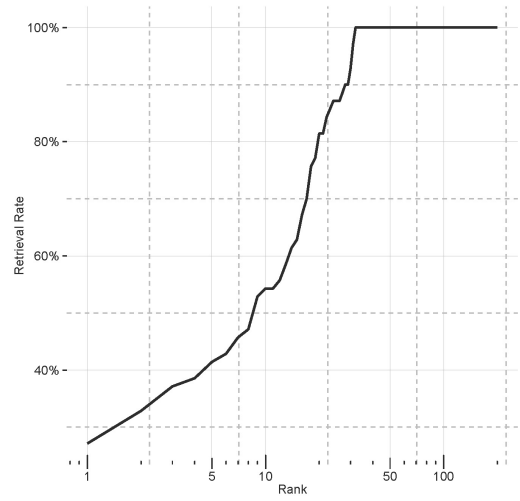


(f) PCA

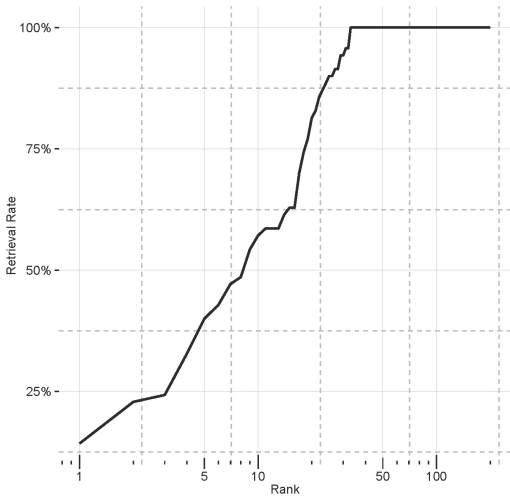
Figure 6.41: Experiment 4 3-5 to 12-14 - Score Histogram



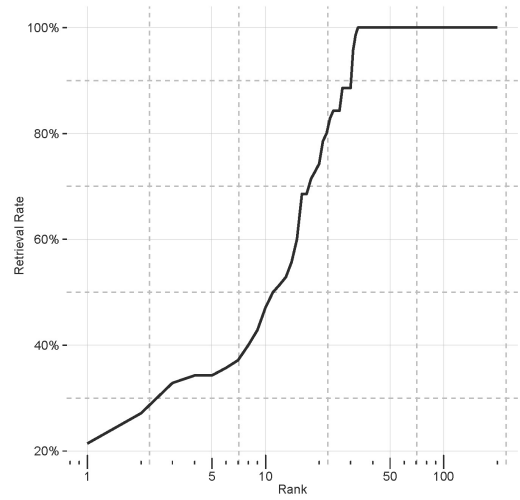
(a) Cognitec



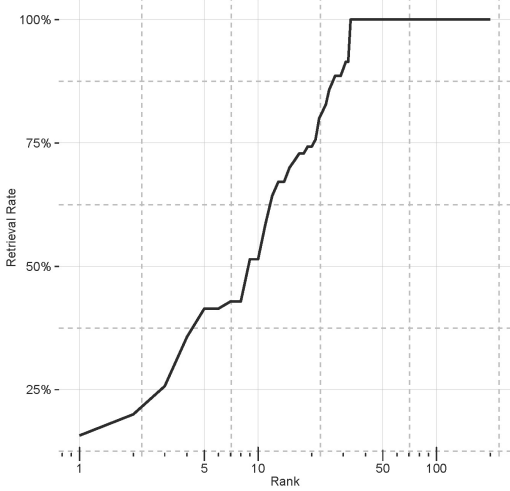
(b) S4F



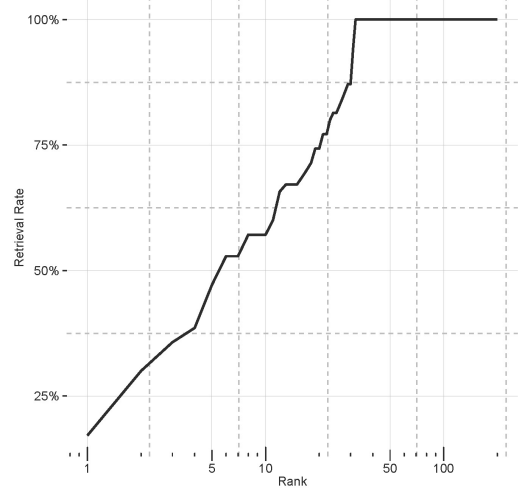
(c) cohortLDA



(d) LRPCA

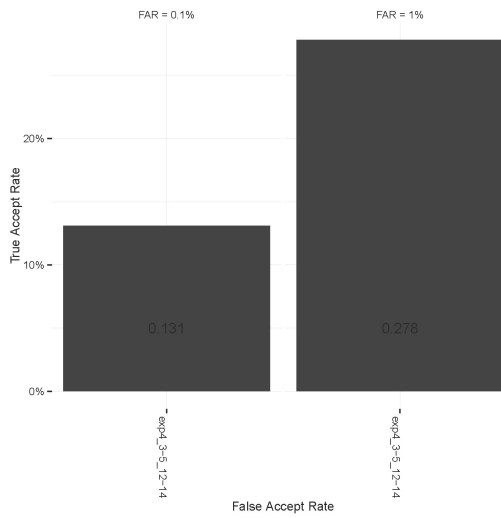


(e) LDA

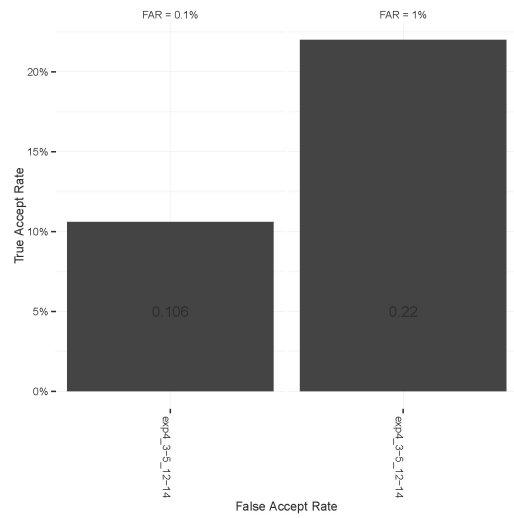


(f) PCA

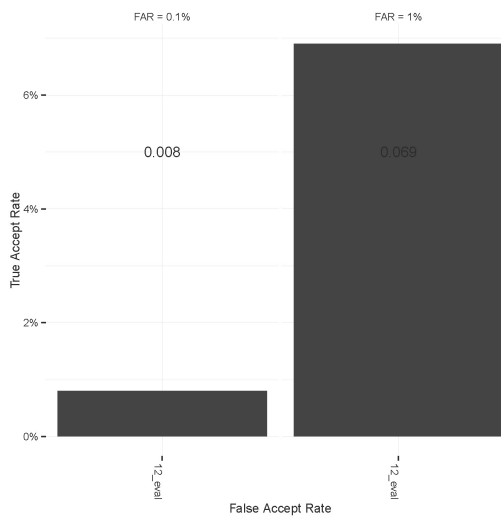
Figure 6.42: Experiment 4 3-5 to 12-14 - CMC



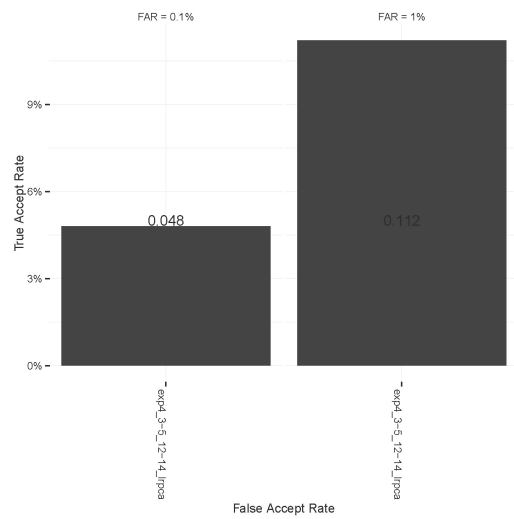
(a) Cognitec



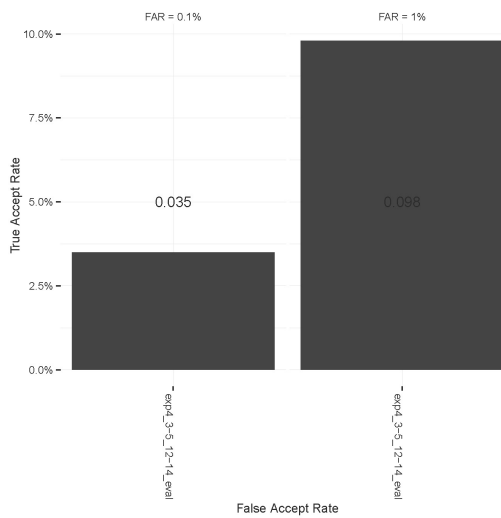
(b) S4F



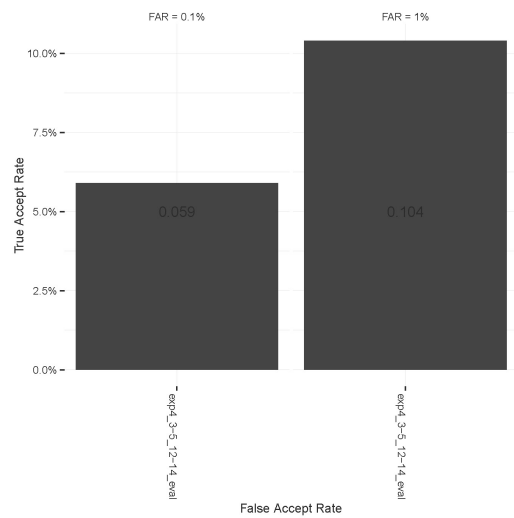
(c) cohortLDA



(d) LRPCA

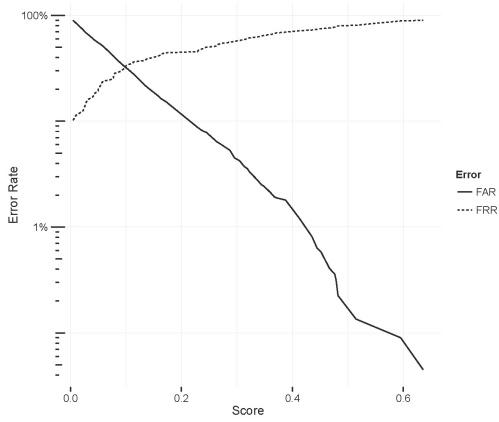


(e) LDA

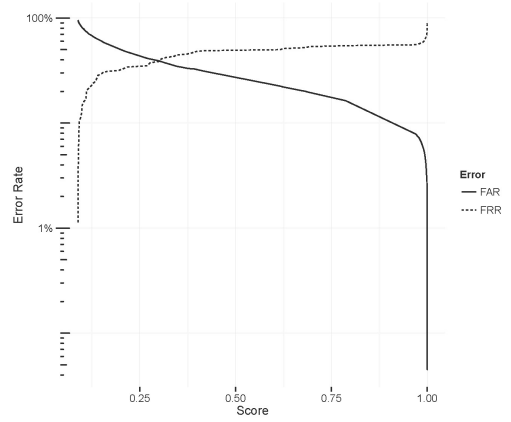


(f) PCA

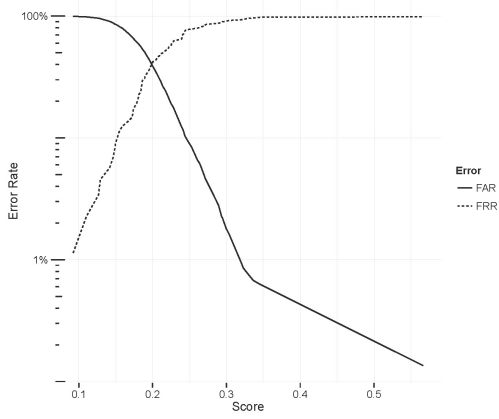
Figure 6.43: Experiment 4 3-5 to 12-14 - ROC Scores



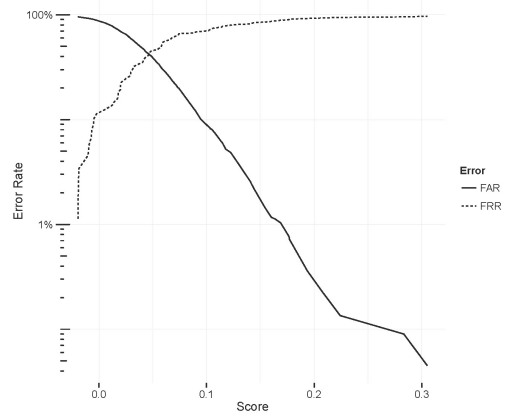
(a) Cognitec



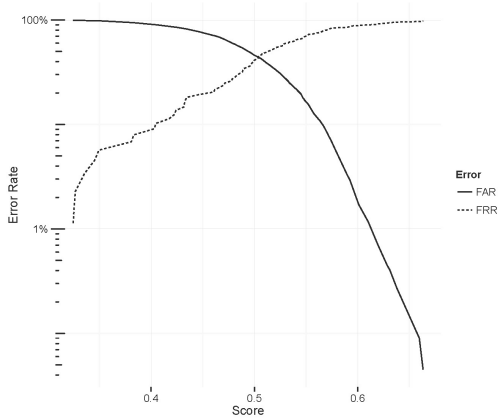
(b) S4F



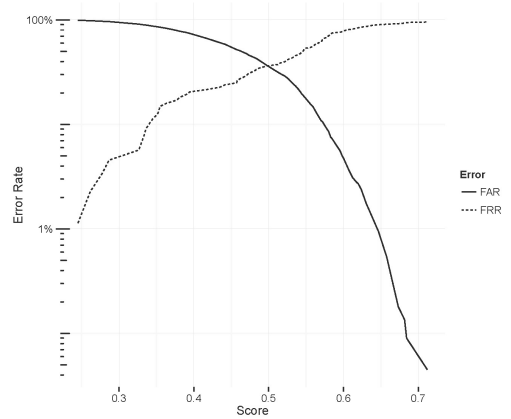
(c) cohortLDA



(d) LRPCA



(e) LDA

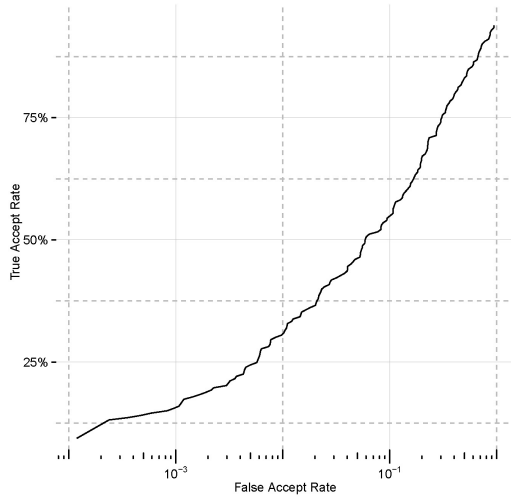


(f) PCA

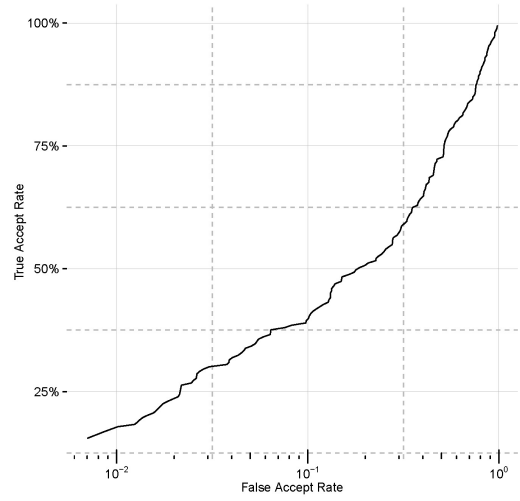
Figure 6.44: Experiment 4 3-5 to 12-14 - EER

APPENDICES F

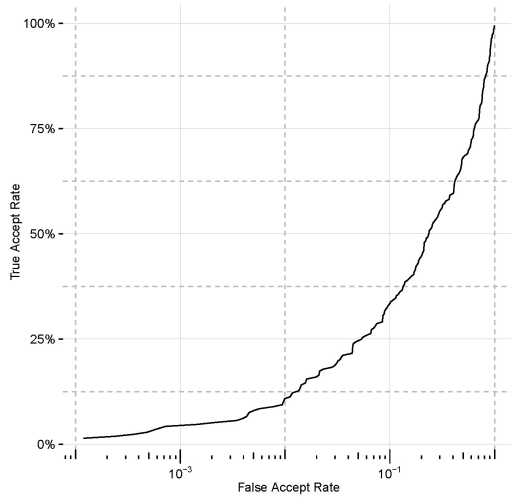
Experiment 4 - 6-8 to 9-11 Results



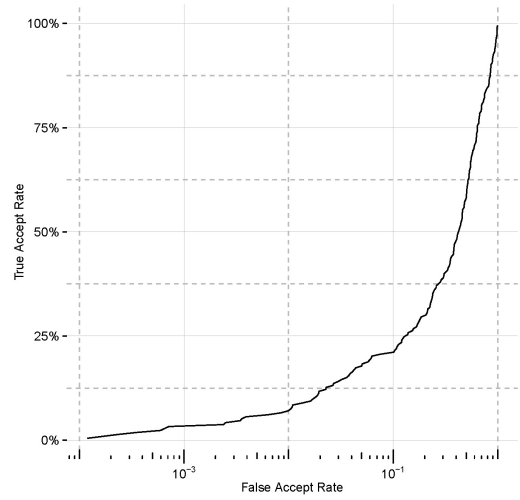
(a) Cognittec



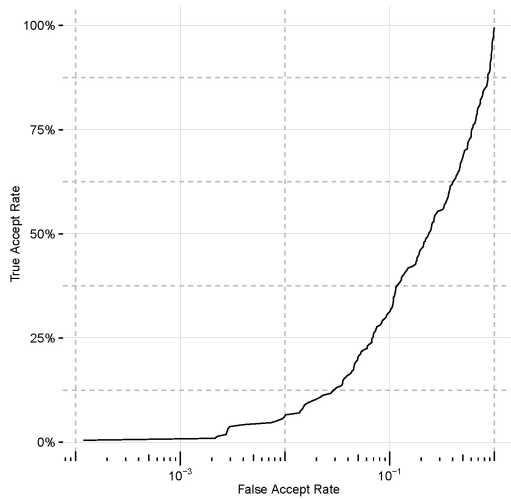
(b) S4F



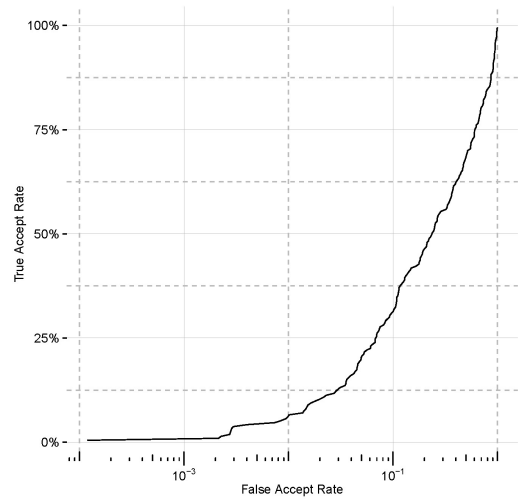
(c) CohortLDA



(d) LRPCA

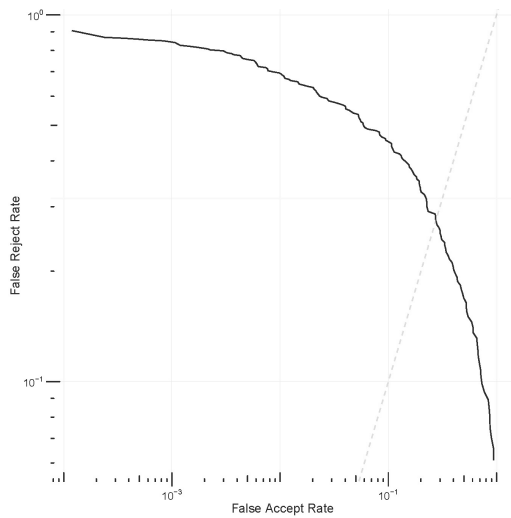


(e) LDA

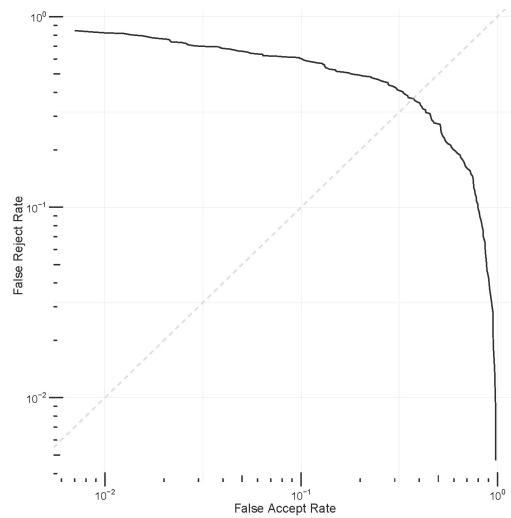


(f) PCA

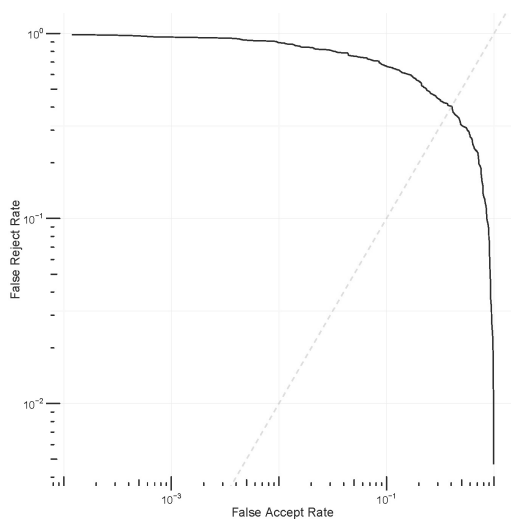
Figure 6.45: Experiment 4 6-8 to 9-11 - ROC



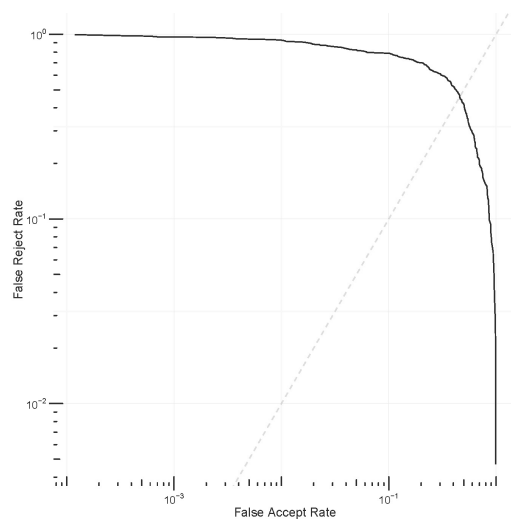
(a) Cognitec



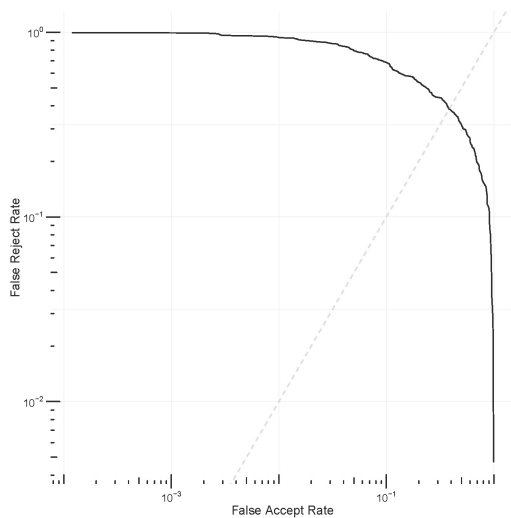
(b) S4F



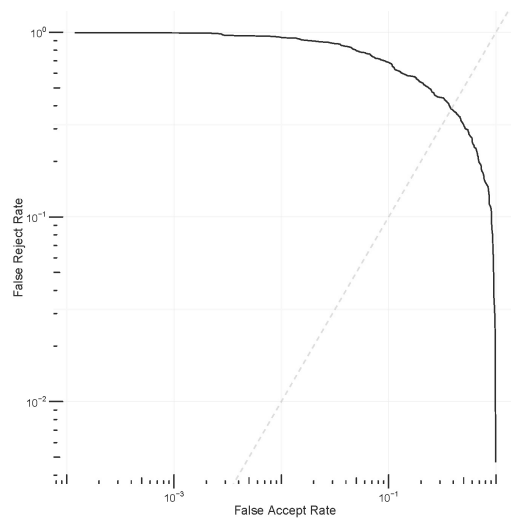
(c) CohortLDA



(d) LRPCA

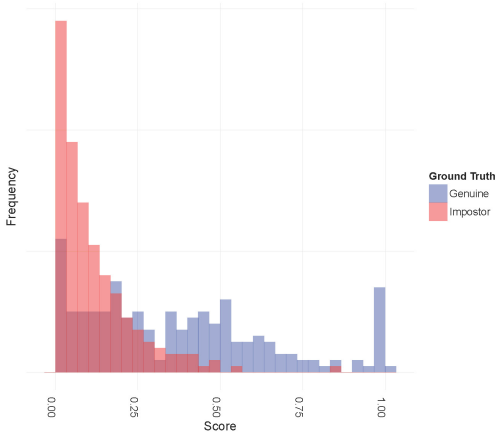


(e) LDA

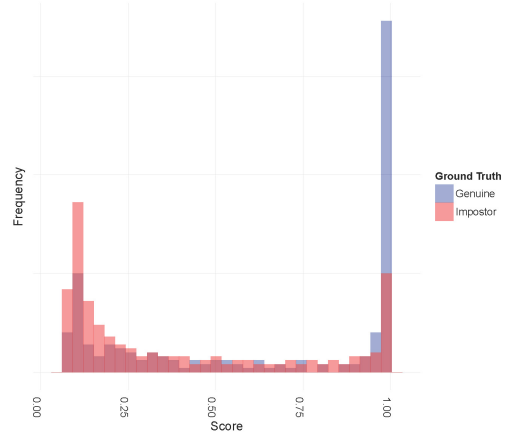


(f) PCA

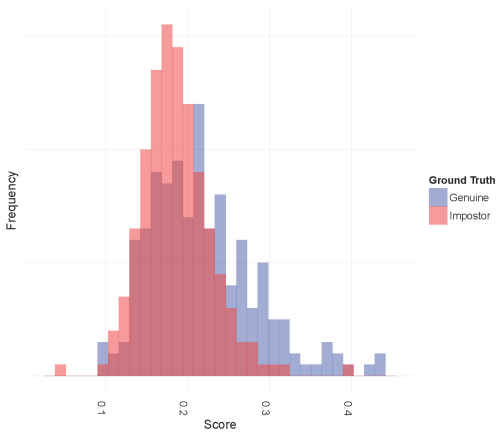
Figure 6.46: Experiment 4 6-8 to 9-11 - DET



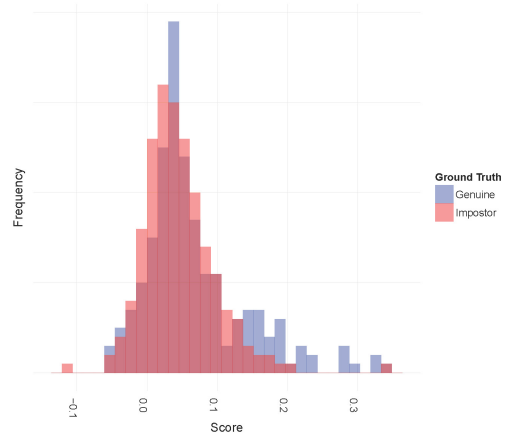
(a) Cognitec



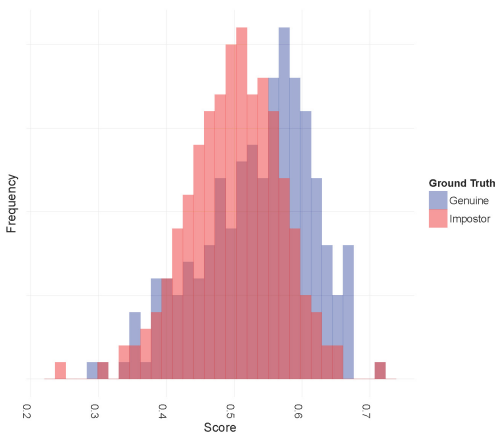
(b) S4F



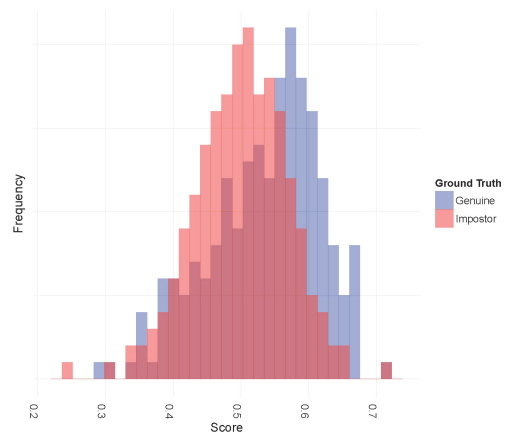
(c) CohortLDA



(d) LRPCA

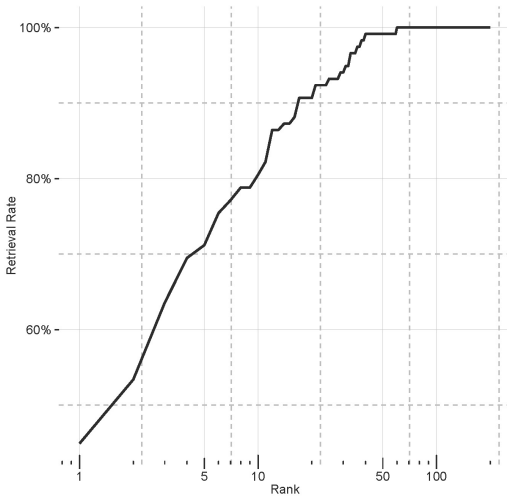


(e) LDA

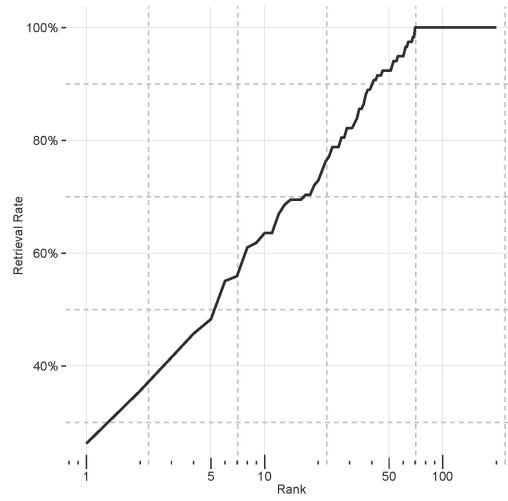


(f) PCA

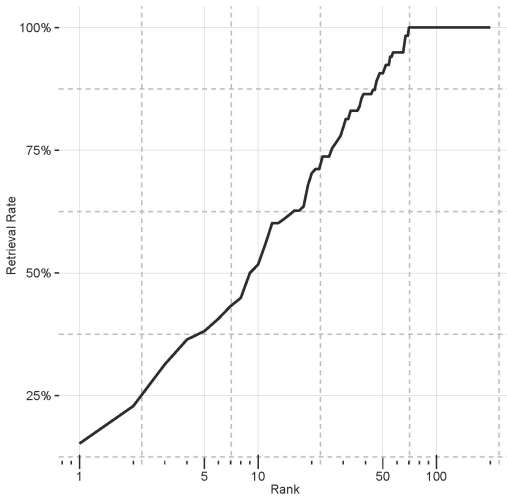
Figure 6.47: Experiment 4 6-8 to 9-11 - Score Histogram



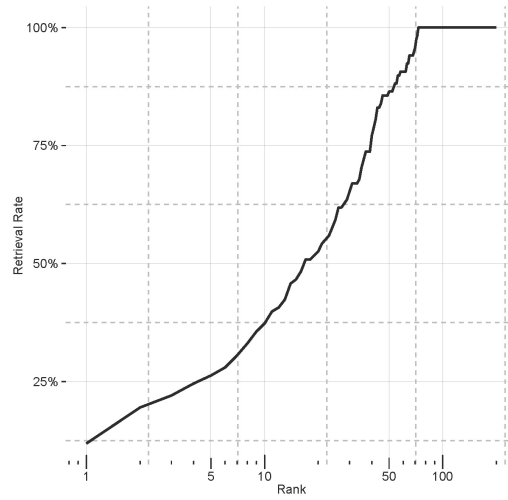
(a) Cognitec



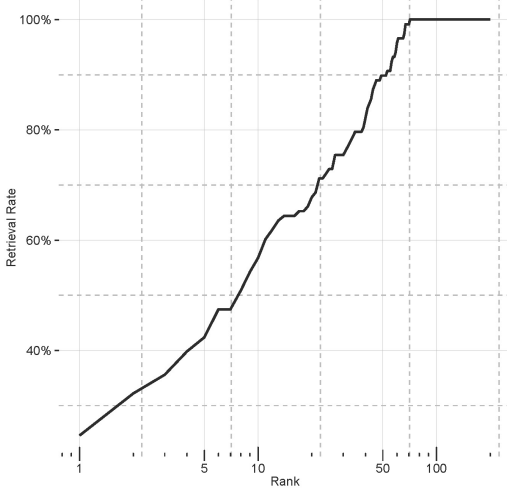
(b) S4F



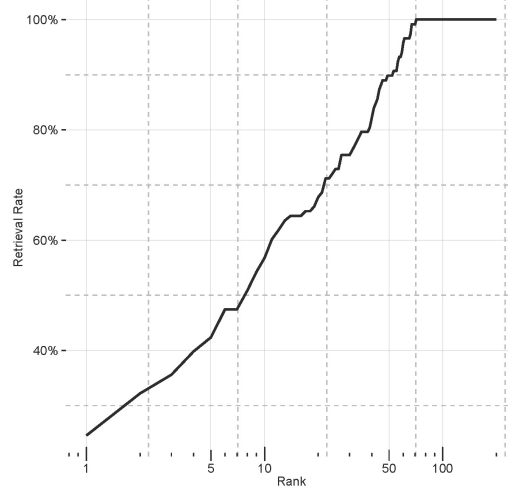
(c) CohortLDA



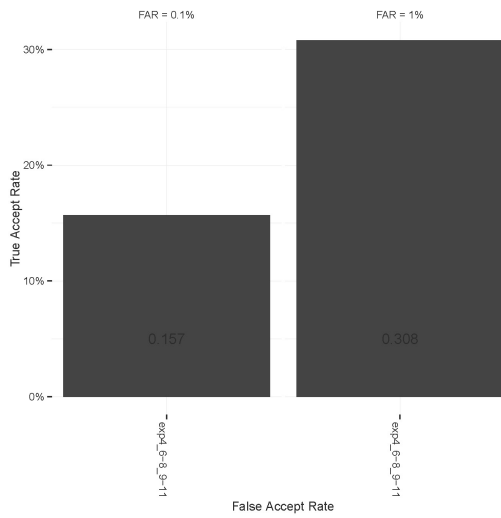
(d) LRPCA



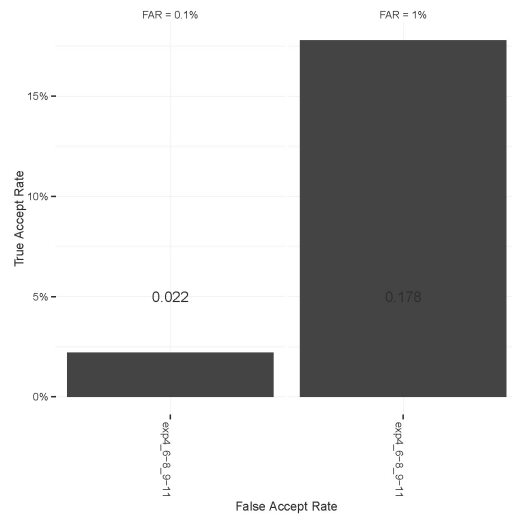
(e) LDA



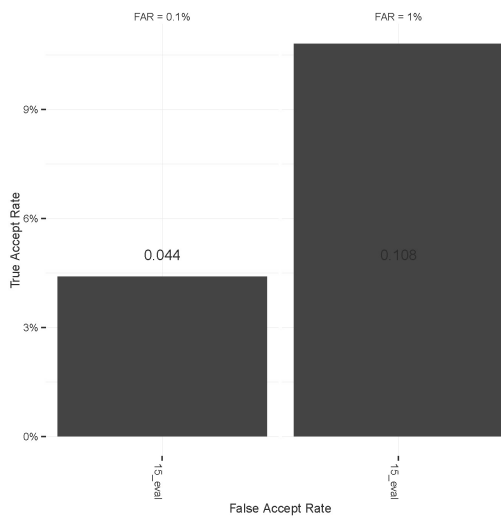
(f) PCA



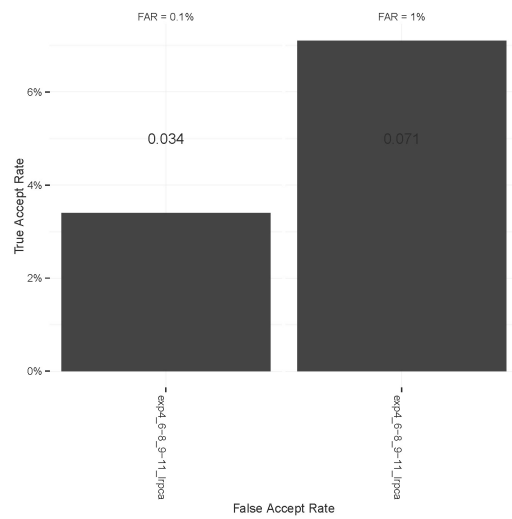
(a) Cognitec



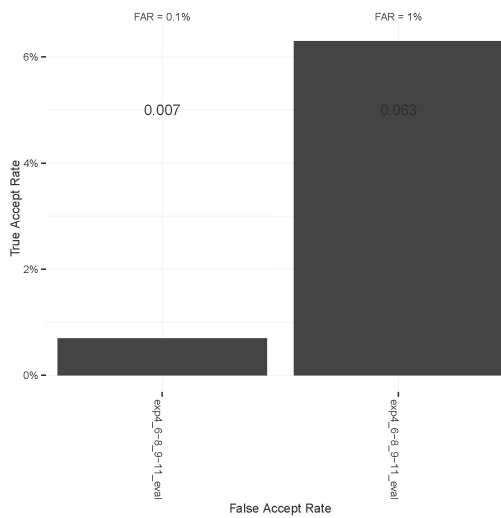
(b) S4F



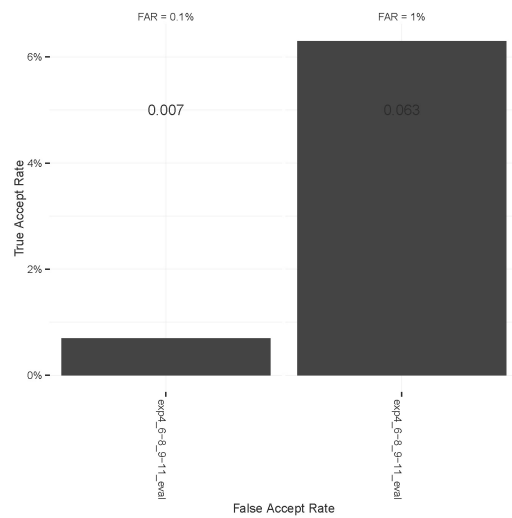
(c) CohortLDA



(d) LRPCA

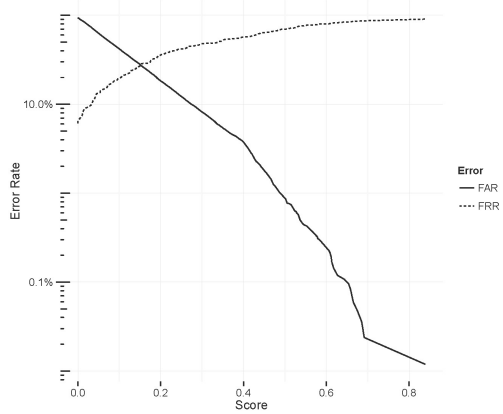


(e) LDA

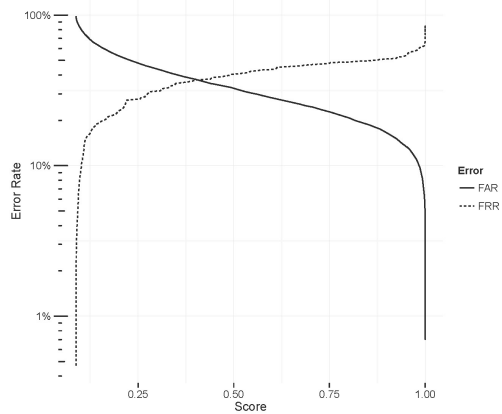


(f) PCA

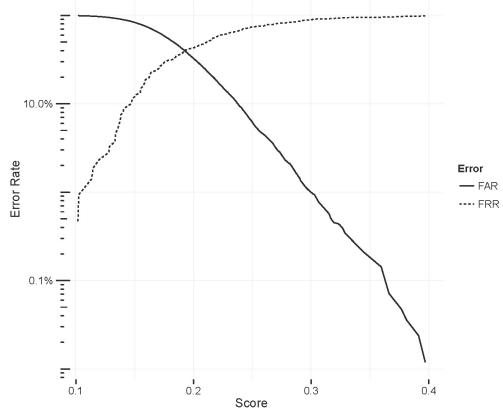
Figure 6.49: Experiment 4 6-8 to 9-11 - ROC Score



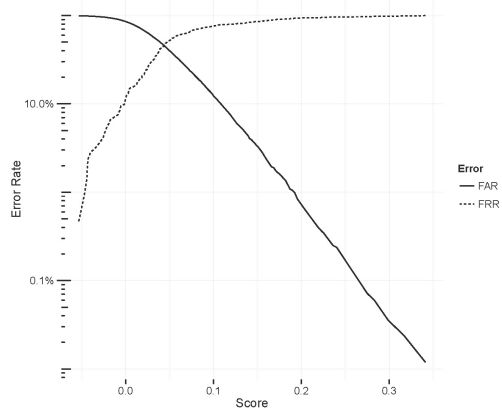
(a) Cognitec



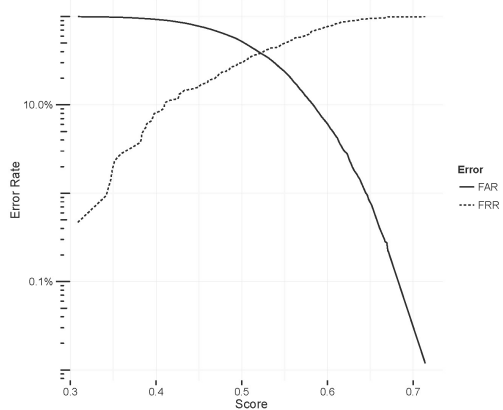
(b) S4F



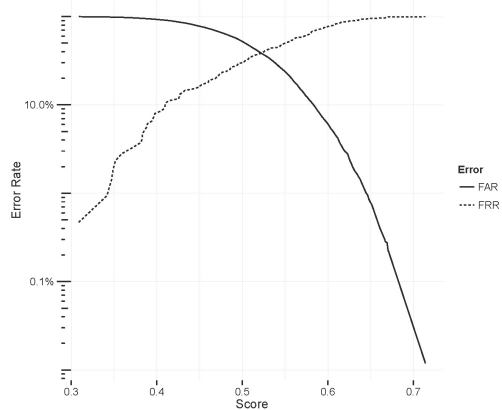
(c) CohortLDA



(d) LRPCA



(e) LDA

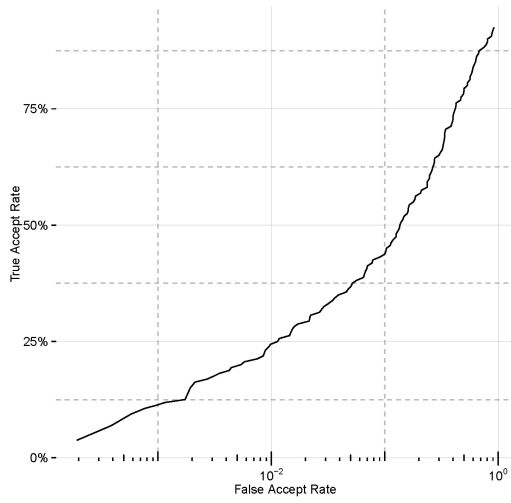


(f) PCA

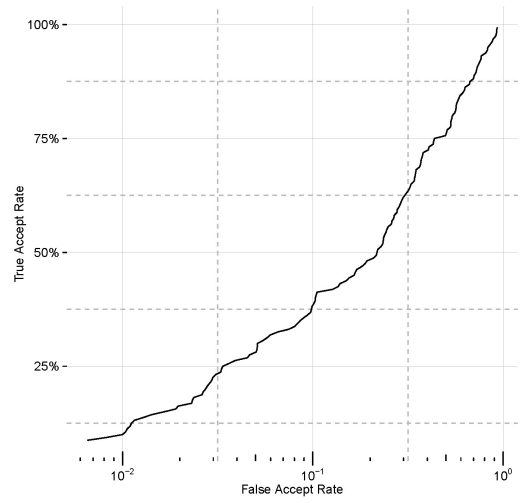
Figure 6.50: Experiment 4 6-8 to 9-11 - EER

APPENDICES G

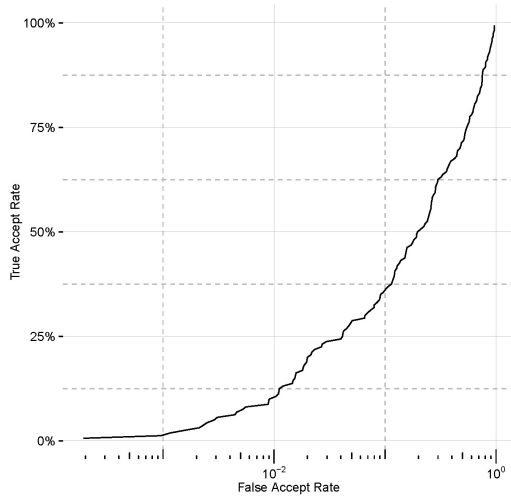
Experiment 4 - 6-8 to 12-14 Results



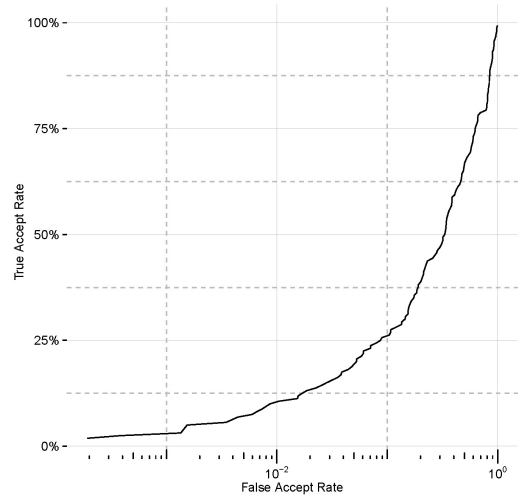
(a) Cognittec



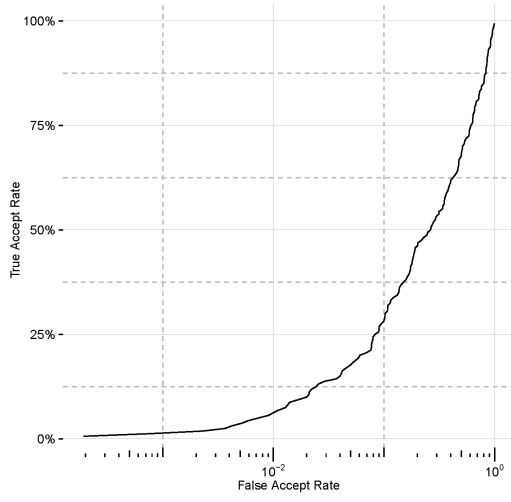
(b) S4F



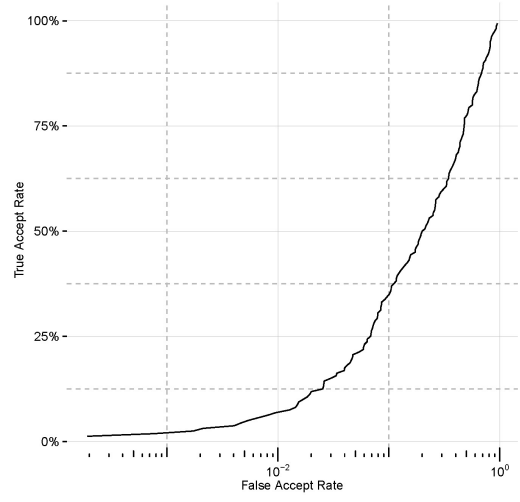
(c) CohortLDA



(d) LRPCA

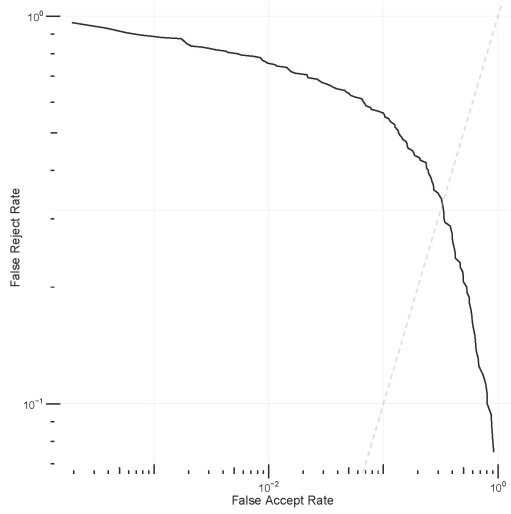


(e) LDA

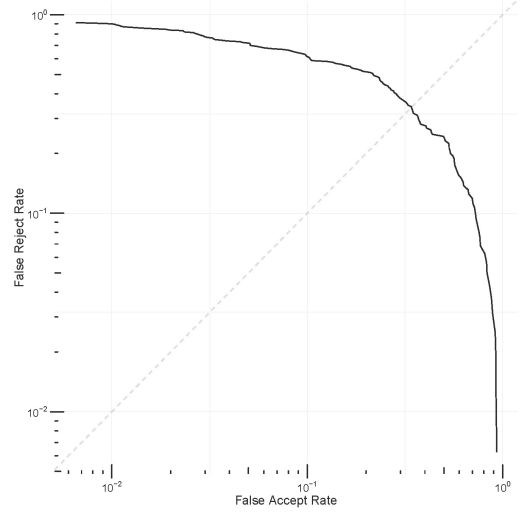


(f) PCA

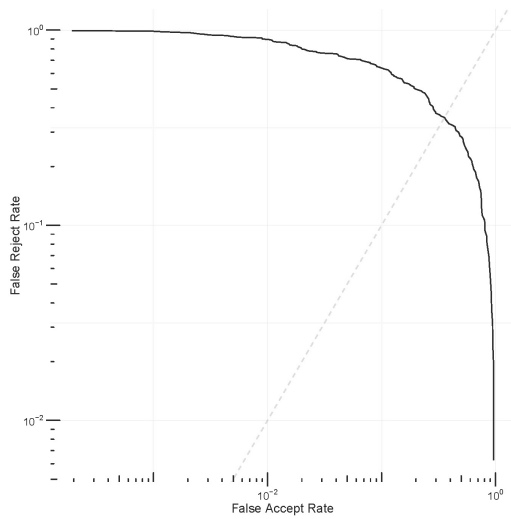
Figure 6.51: Experiment 4 6-8 to 12-14 - ROC



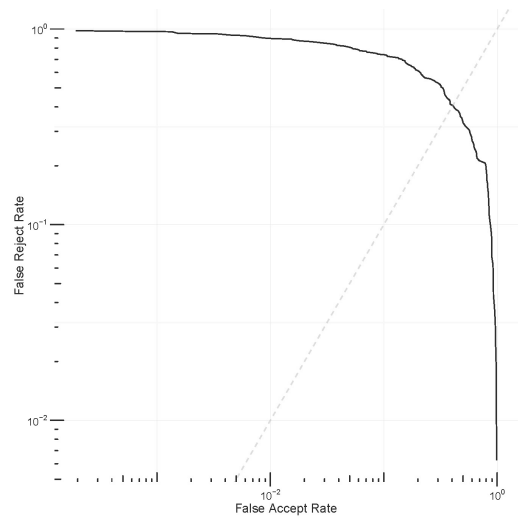
(a) Cognitec



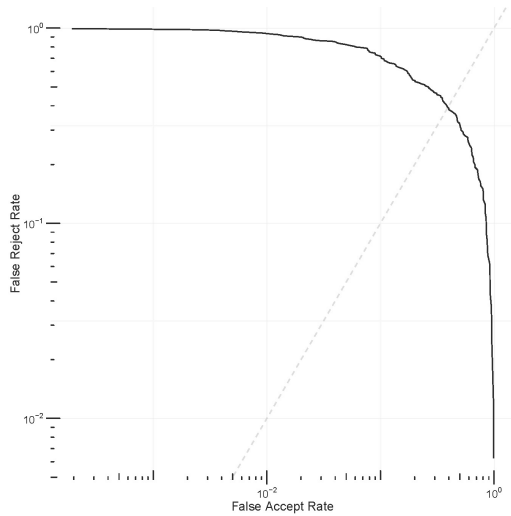
(b) S4F



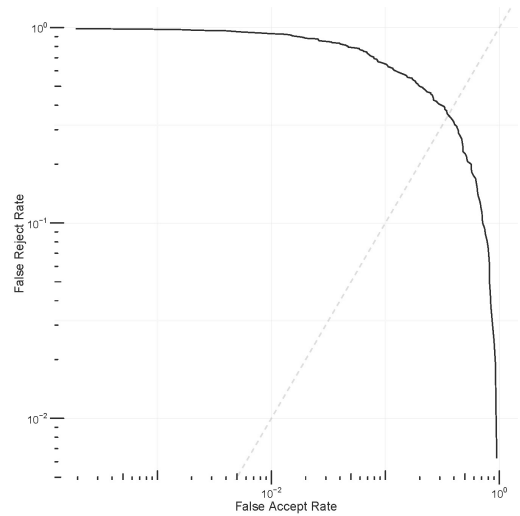
(c) CohortLDA



(d) LRPCA

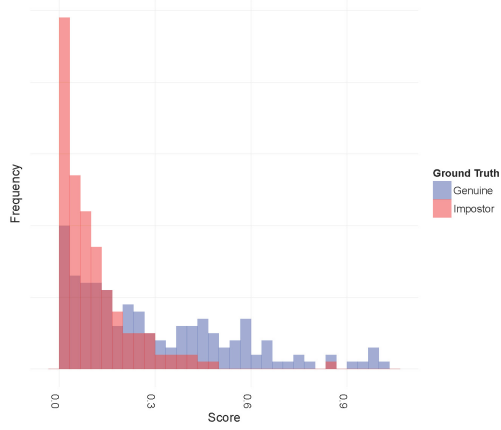


(e) LDA

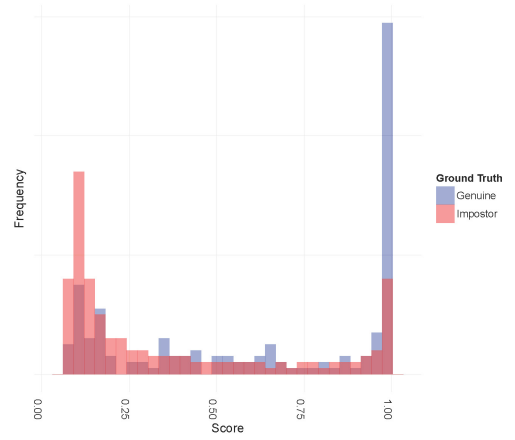


(f) PCA

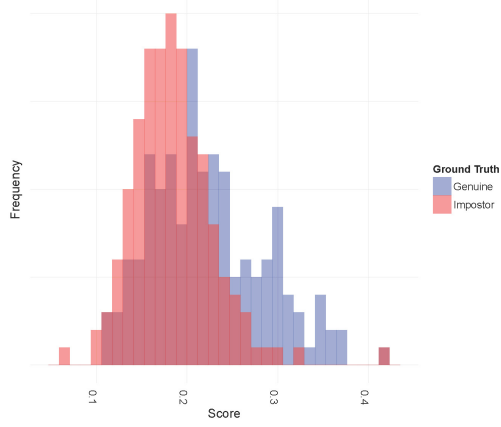
Figure 6.52: Experiment 4 6-8 to 12-14 - DET



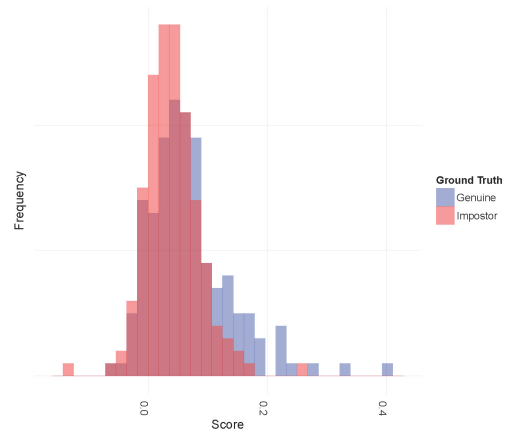
(a) Cognitec



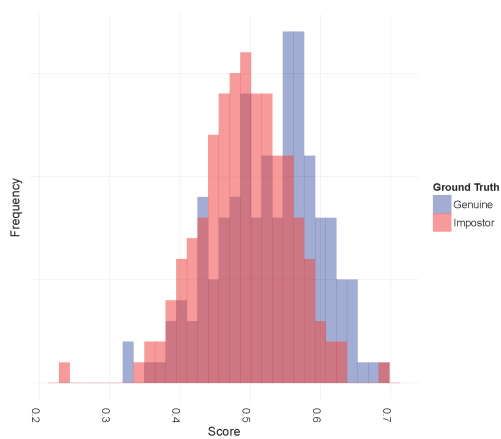
(b) S4F



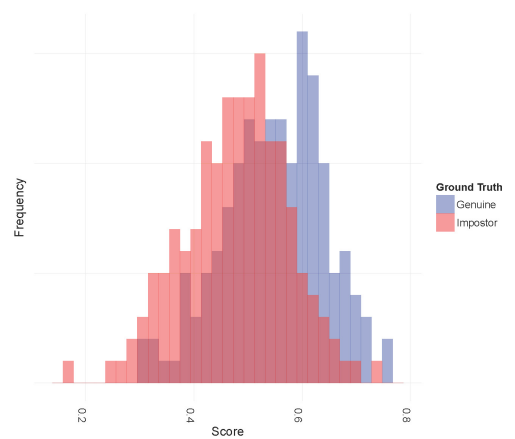
(c) CohortLDA



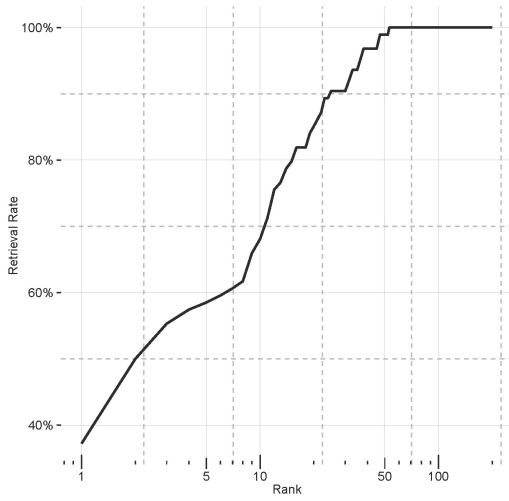
(d) LRPCA



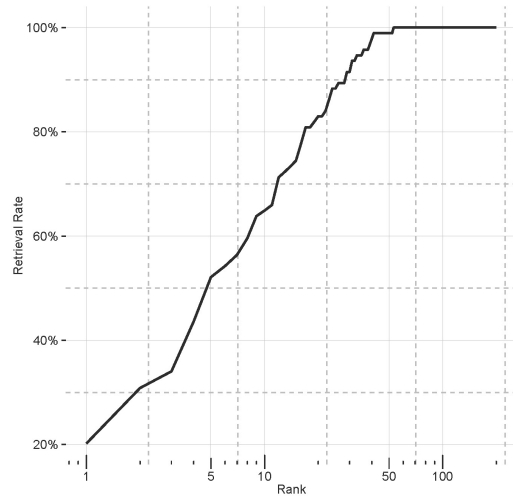
(e) LDA



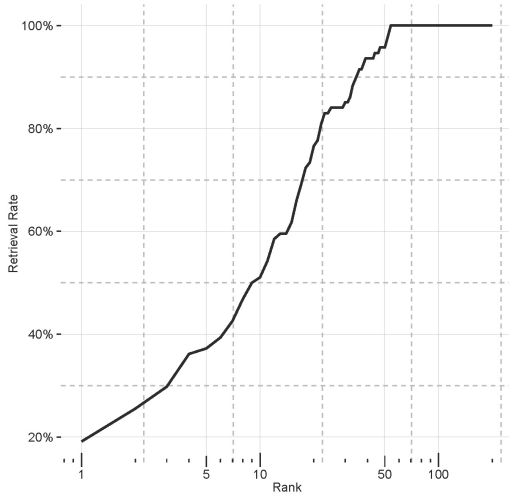
(f) PCA



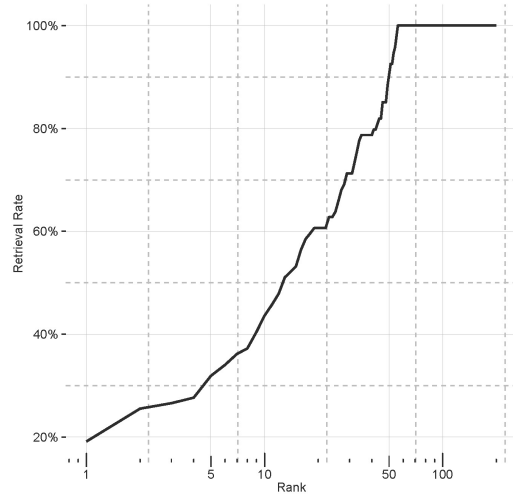
(a) Cognitec



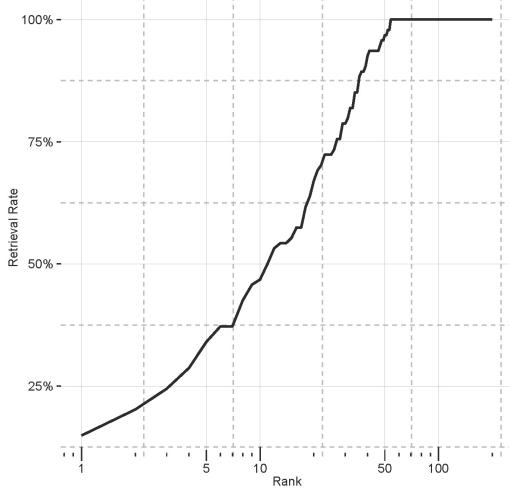
(b) S4F



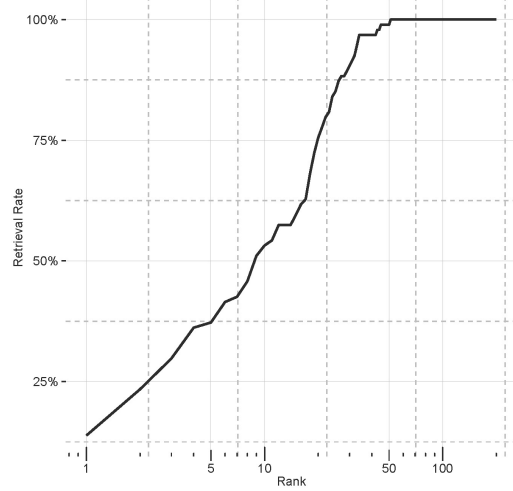
(c) CohortLDA



(d) LRPCA

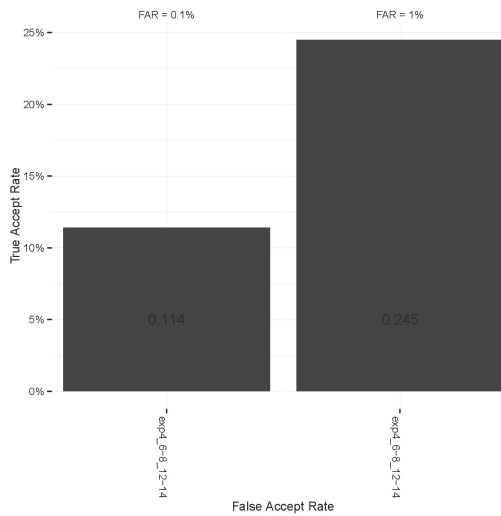


(e) LDA

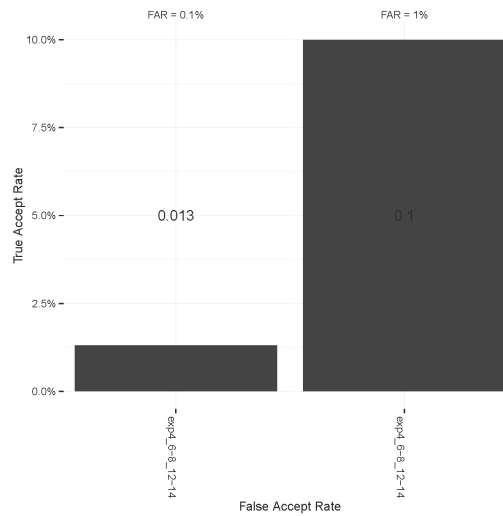


(f) PCA

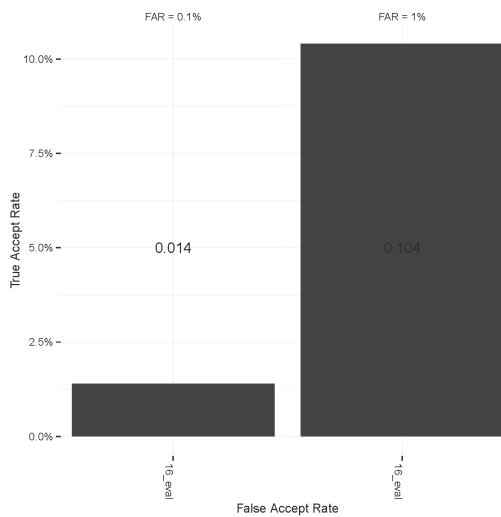
Figure 6.54: Experiment 4 6-8 to 12-14 - CMC



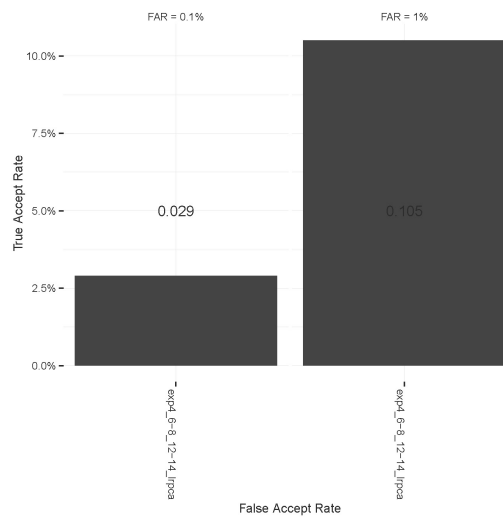
(a) Cognitec



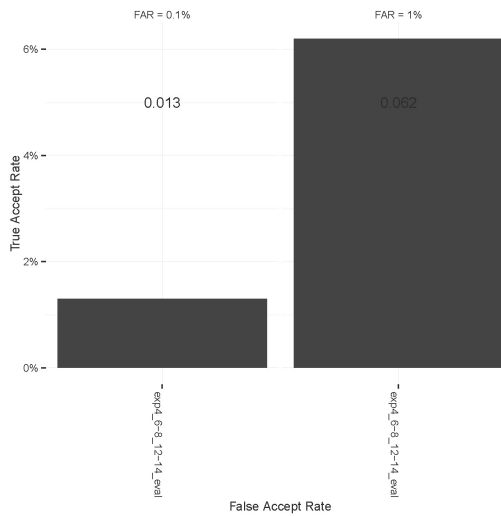
(b) S4F



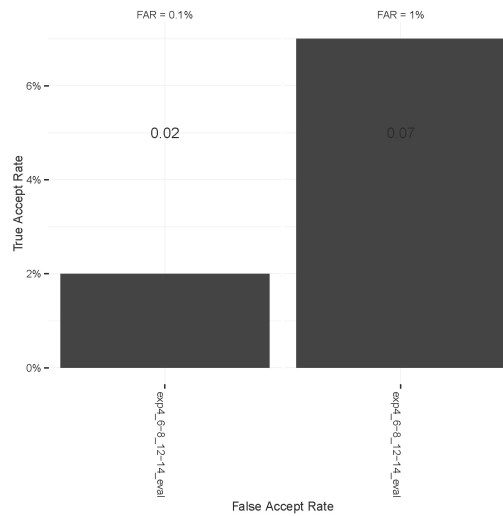
(c) CohortLDA



(d) LRPCA

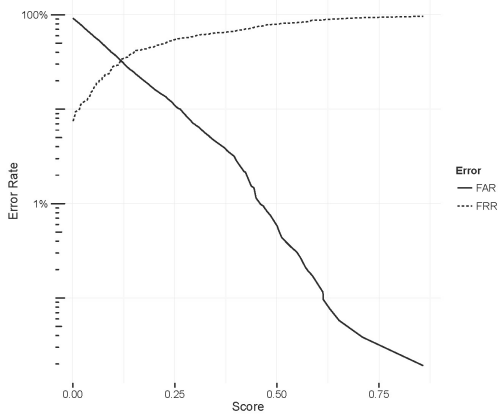


(e) LDA

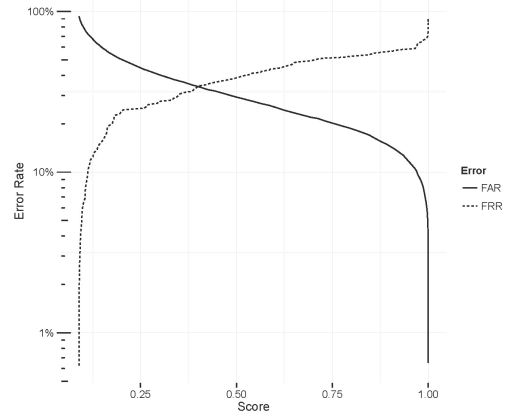


(f) PCA

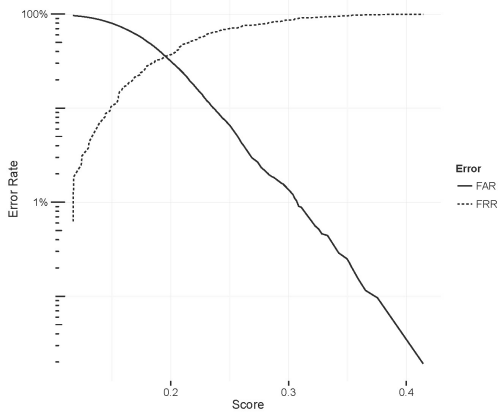
Figure 6.55: Experiment 4 6-8 to 12-14 - ROC Scores



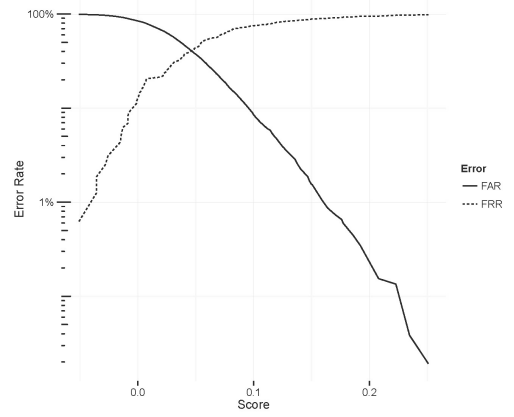
(a) Cognitec



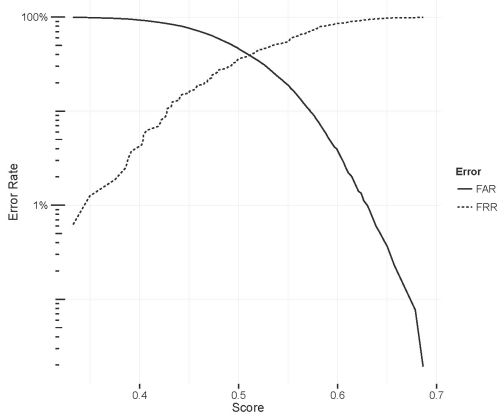
(b) S4F



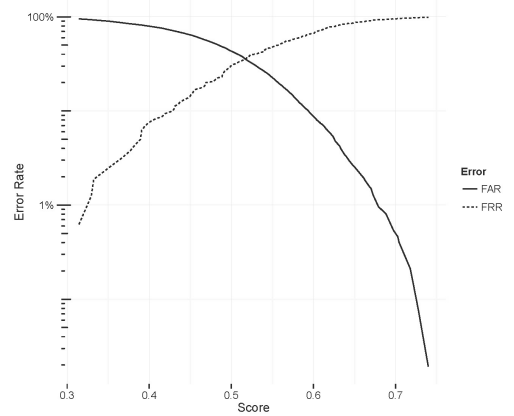
(c) CohortLDA



(d) LRPCA



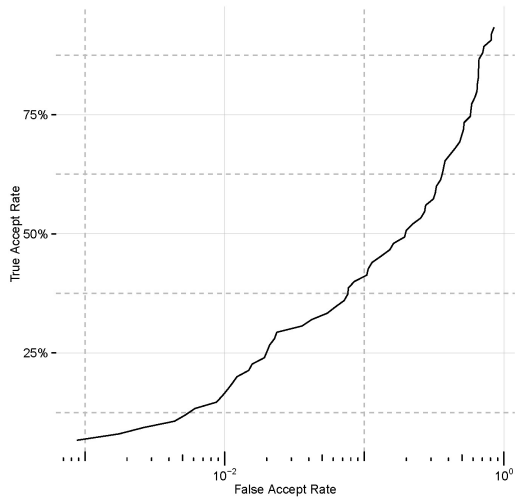
(e) LDA



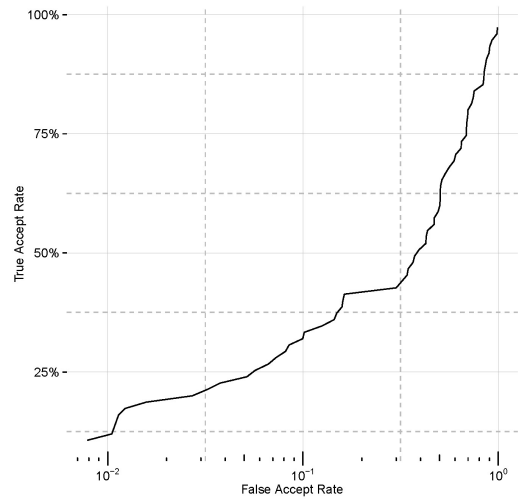
(f) PCA

APPENDICES H

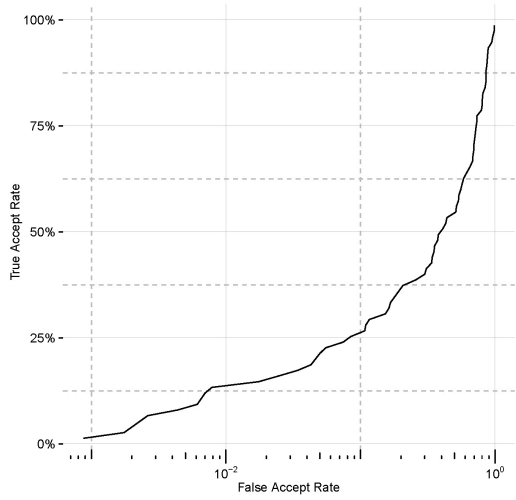
Experiment 4 - 6-8 to 15-17 Results



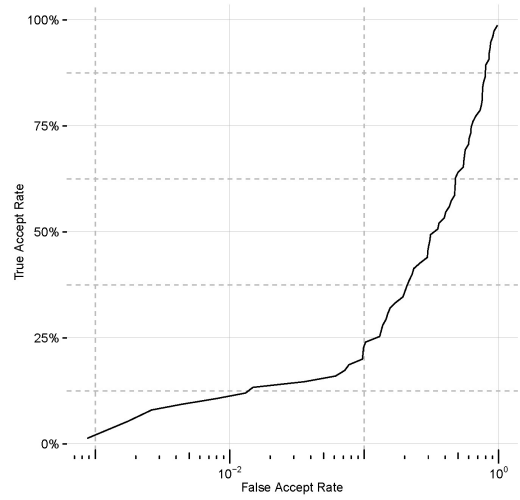
(a) Cognitec



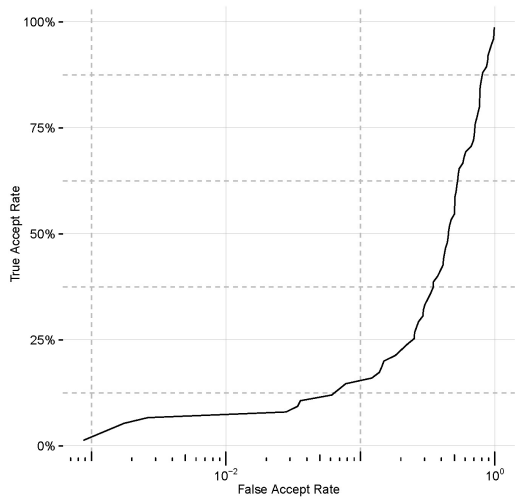
(b) S4F



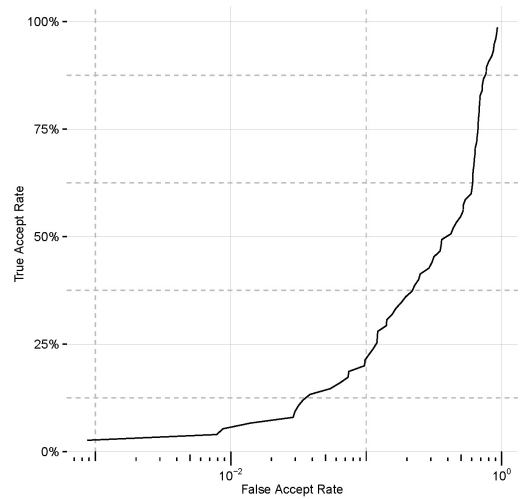
(c) CohortLDA



(d) LRPCA

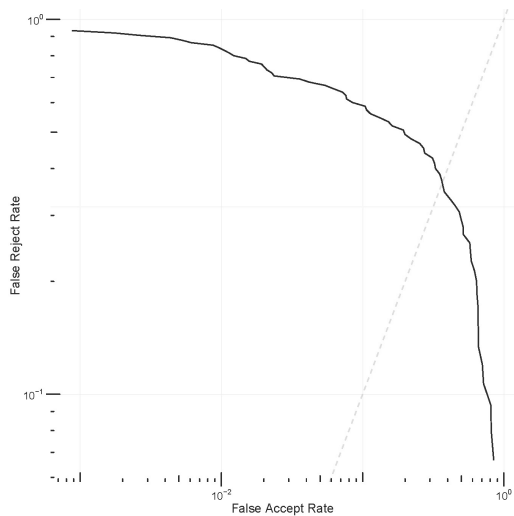


(e) LDA

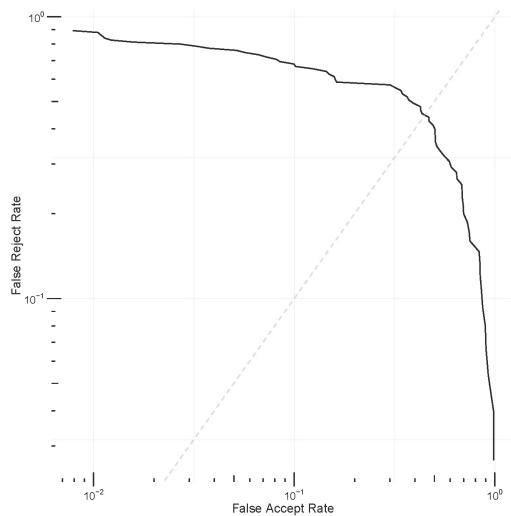


(f) PCA

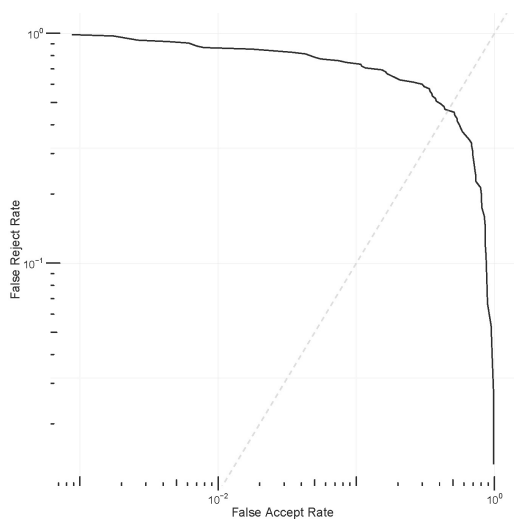
Figure 6.57: Experiment 4 6-8 to 15-17 - ROC



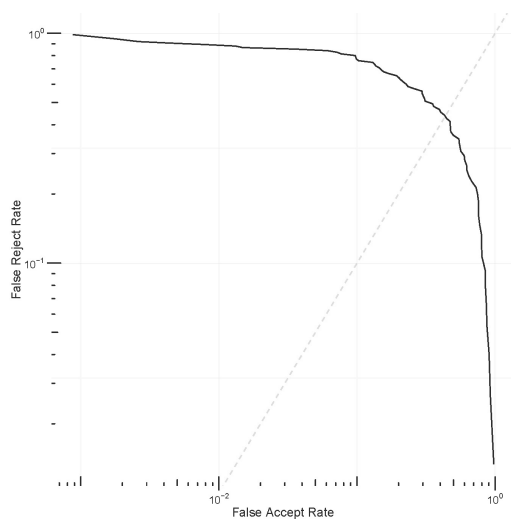
(a) Cognitec



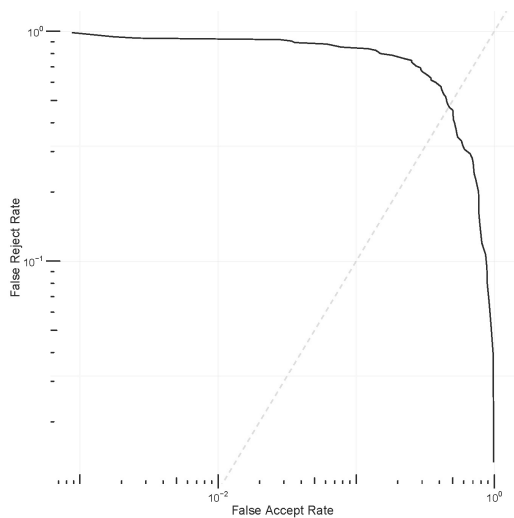
(b) S4F



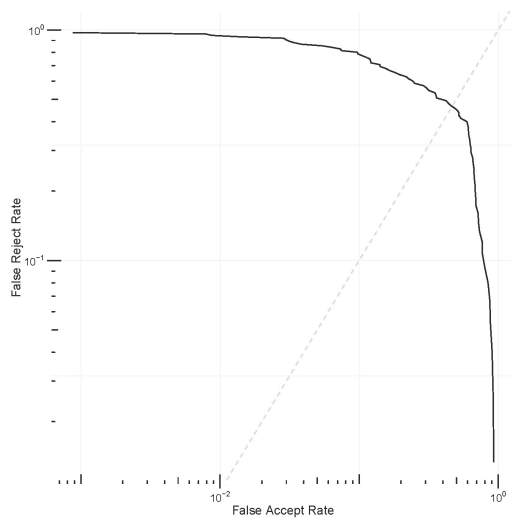
(c) CohortLDA



(d) LRPCA

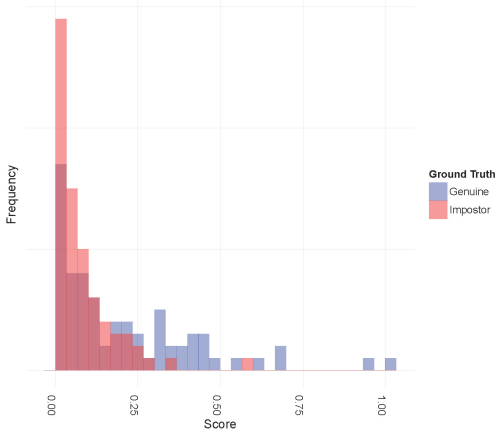


(e) LDA

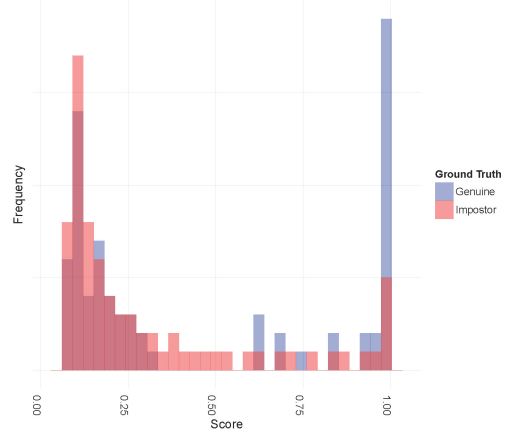


(f) PCA

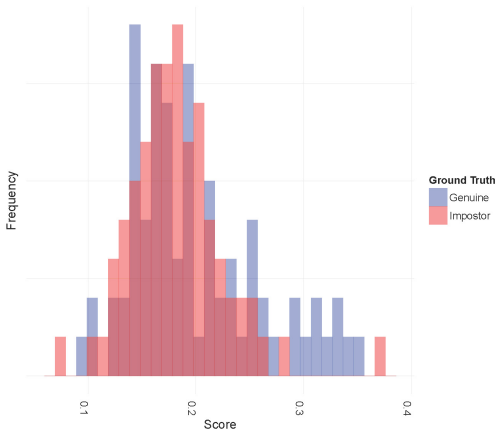
Figure 6.58: Experiment 4 6-8 to 15-17 - DET



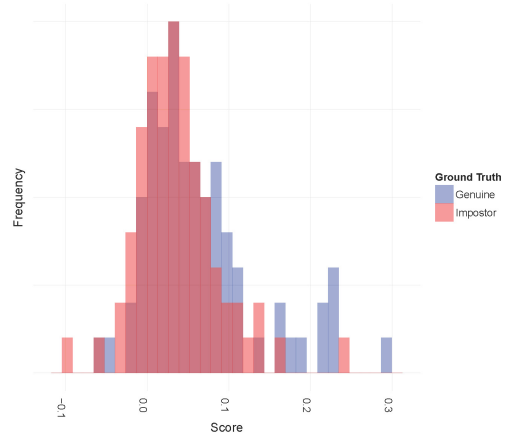
(a) Cognitec



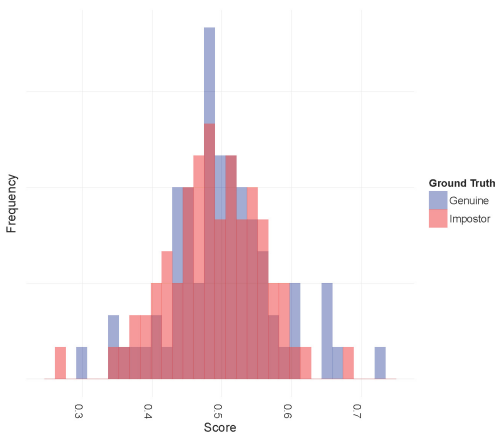
(b) S4F



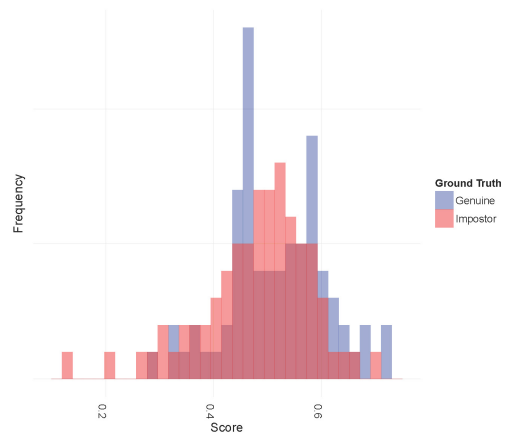
(c) CohortLDA



(d) LRPCA

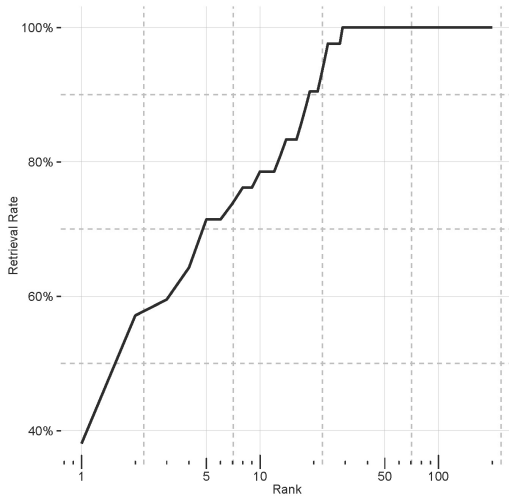


(e) LDA

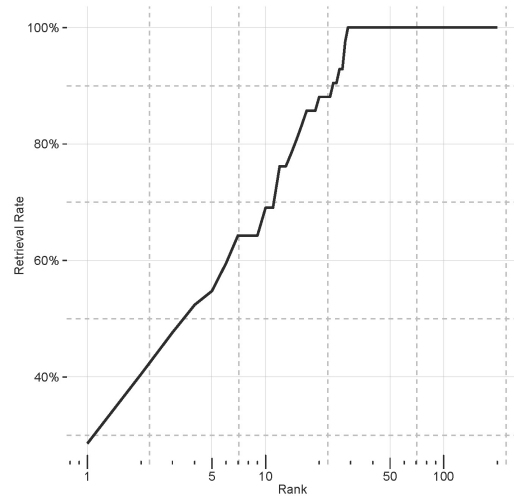


(f) PCA

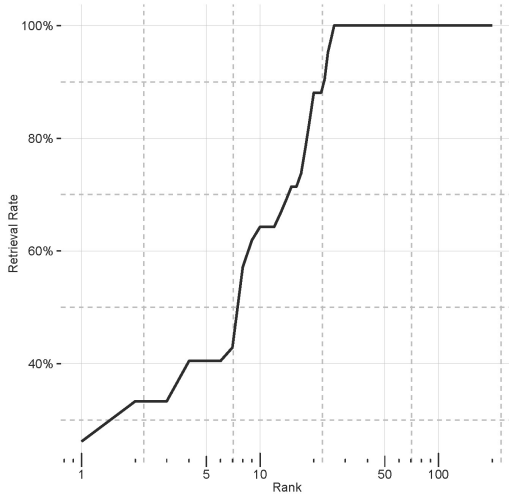
Figure 6.59: Experiment 4 6-8 to 15-17 - Score Histogram



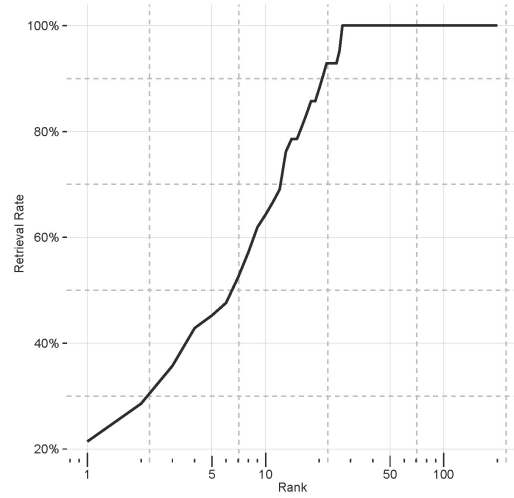
(a) Cognitec



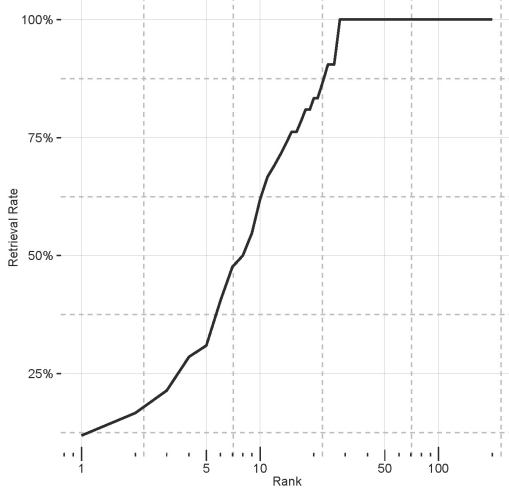
(b) S4F



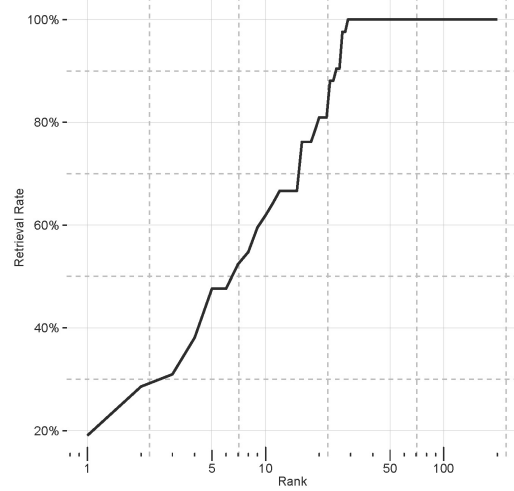
(c) CohortLDA



(d) LRPCA

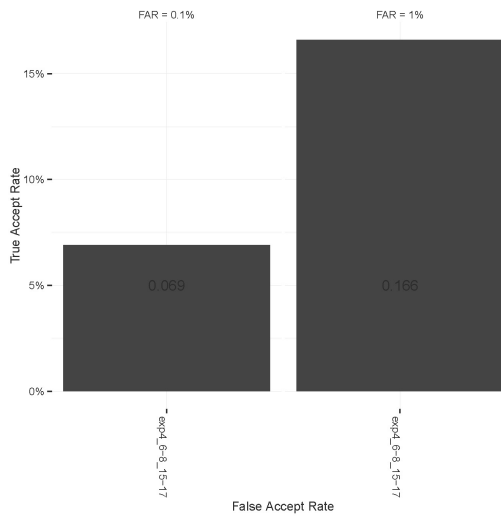


(e) LDA

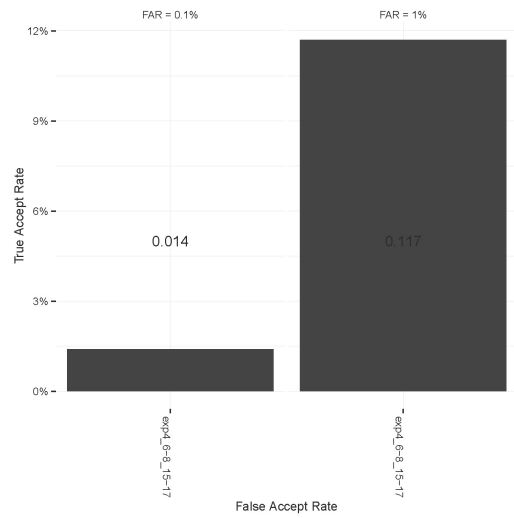


(f) PCA

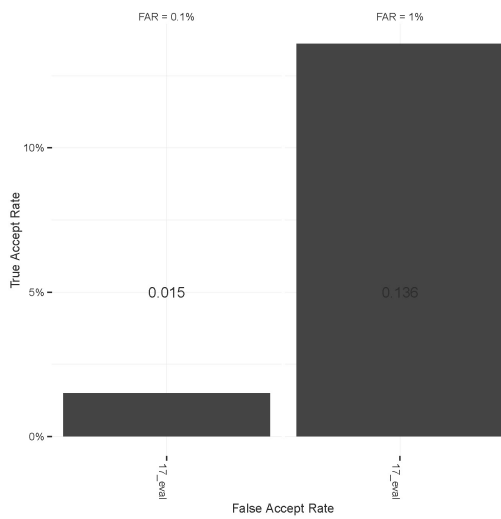
Figure 6.60: Experiment 4 6-8 to 15-17 - CMC



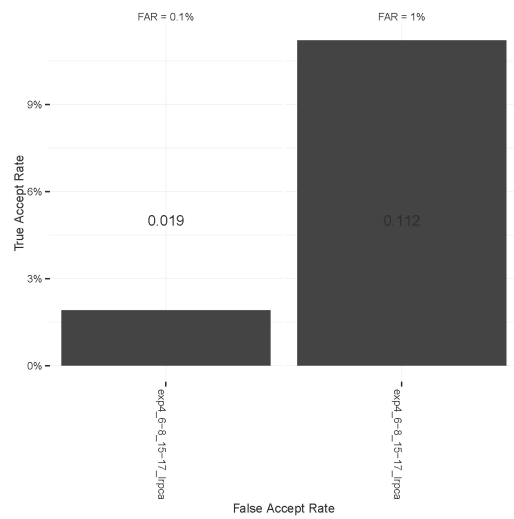
(a) Cognitec



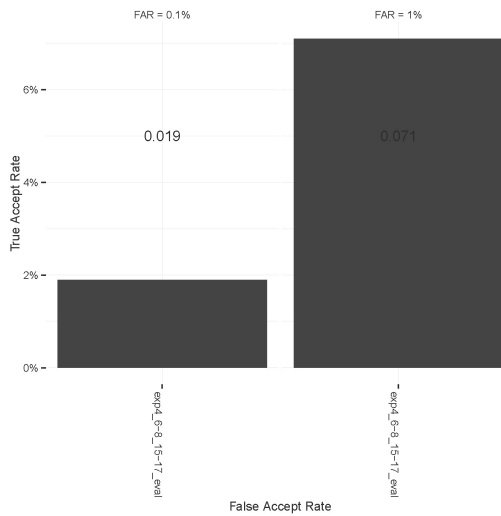
(b) S4F



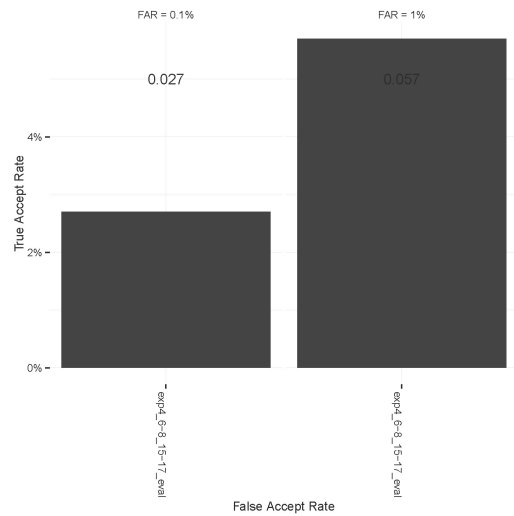
(c) CohortLDA



(d) LRPCA

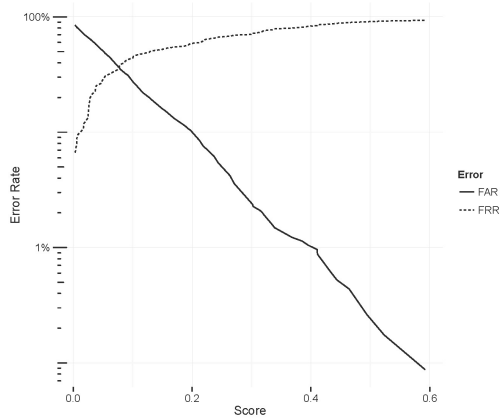


(e) LDA

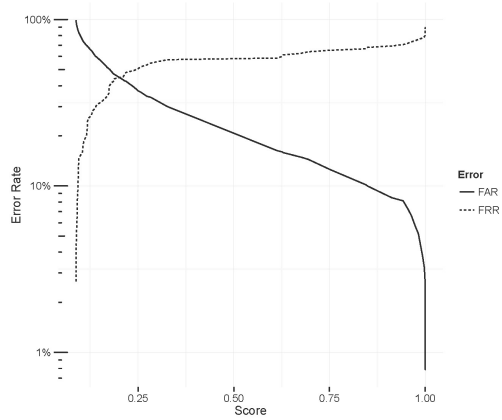


(f) PCA

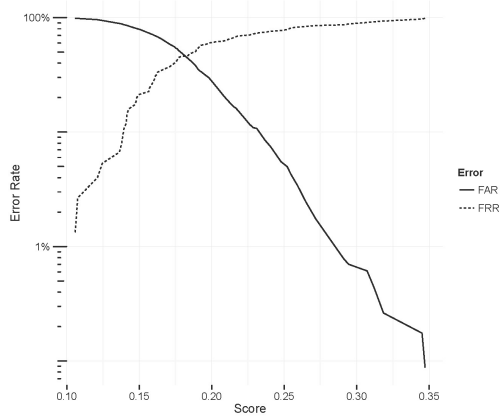
Figure 6.61: Experiment 4 6-8 to 15-17 - ROC Scores



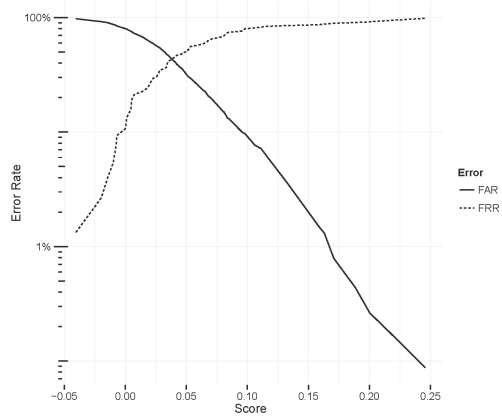
(a) Cognitec



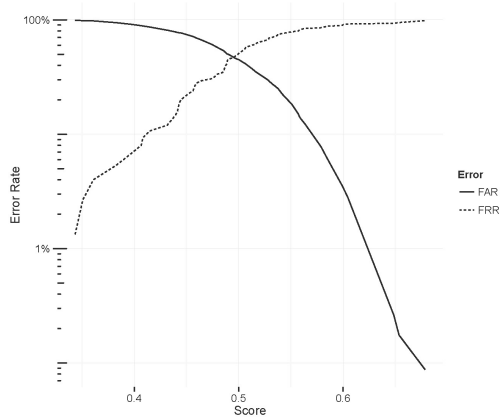
(b) S4F



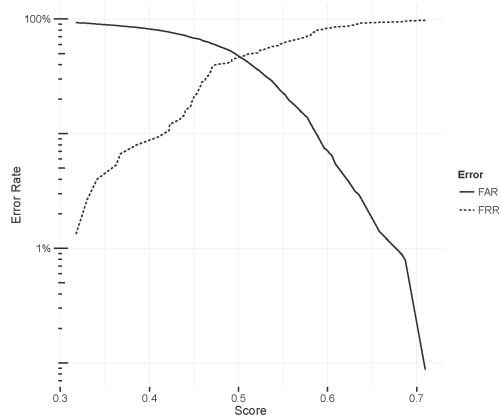
(c) CohortLDA



(d) LRPCA



(e) LDA

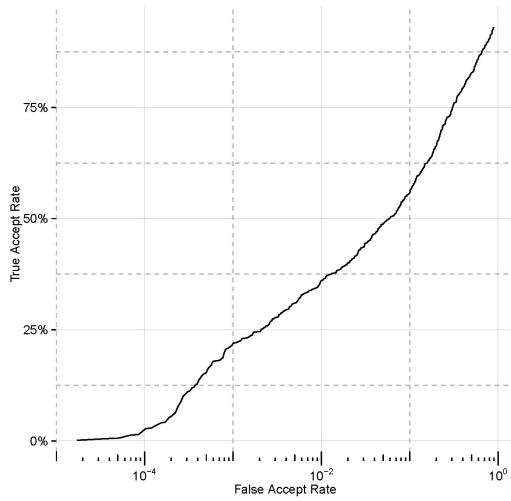


(f) PCA

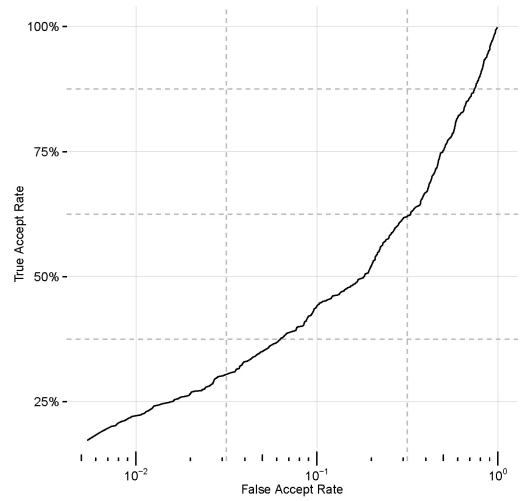
Figure 6.62: Experiment 4 6-8 to 15-17 - EER

APPENDICES I

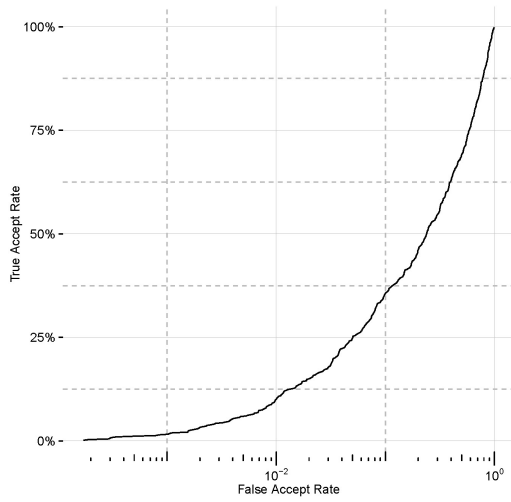
Experiment 4 - 9-11 to 12-14 Results



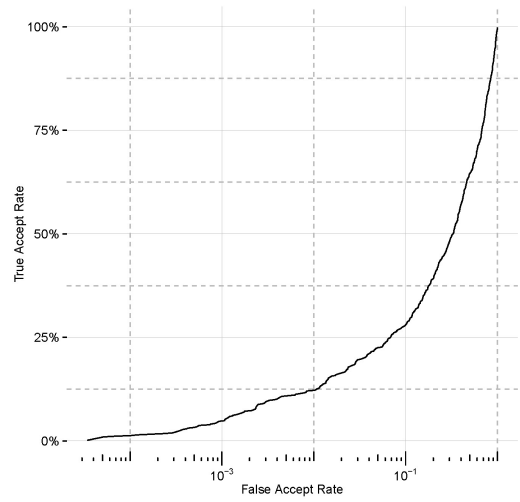
(a) Cognitec



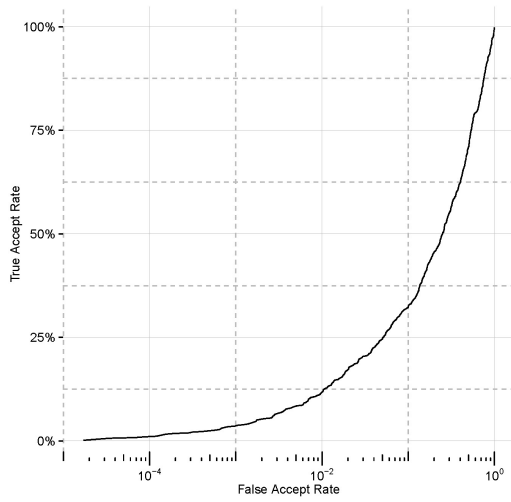
(b) S4F



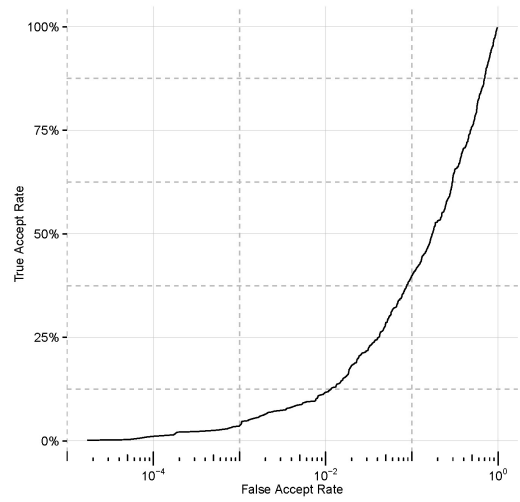
(c) CohortLDA



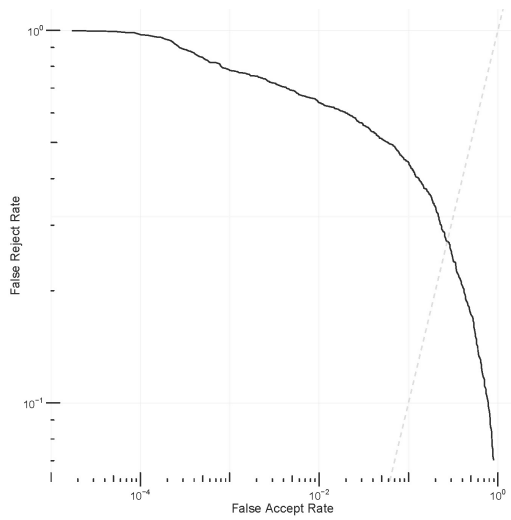
(d) LRPCA



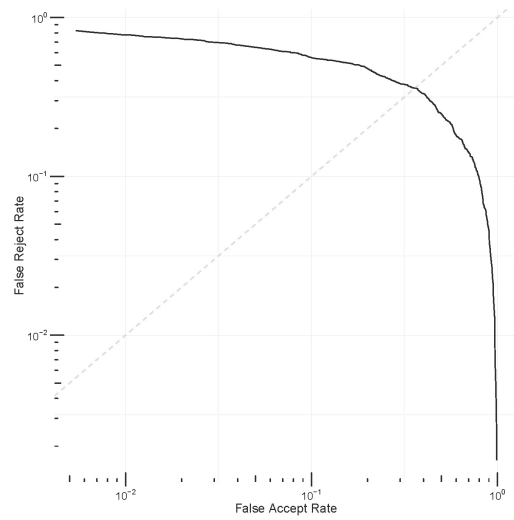
(e) LDA



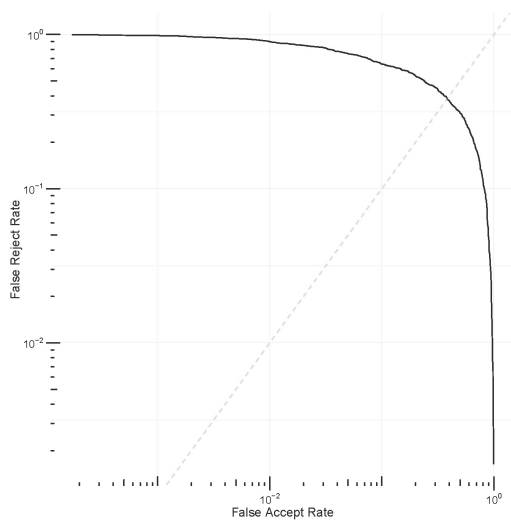
(f) PCA



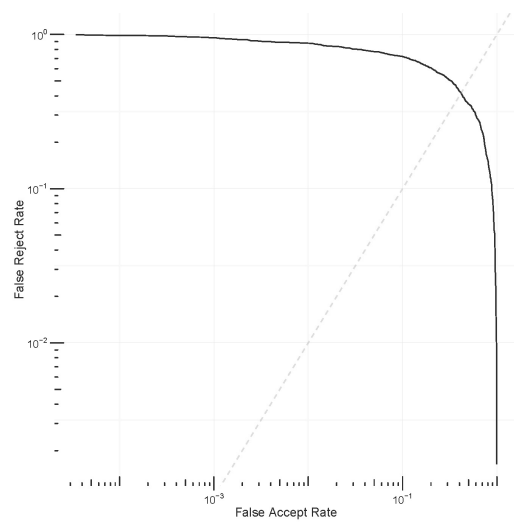
(a) Cognitec



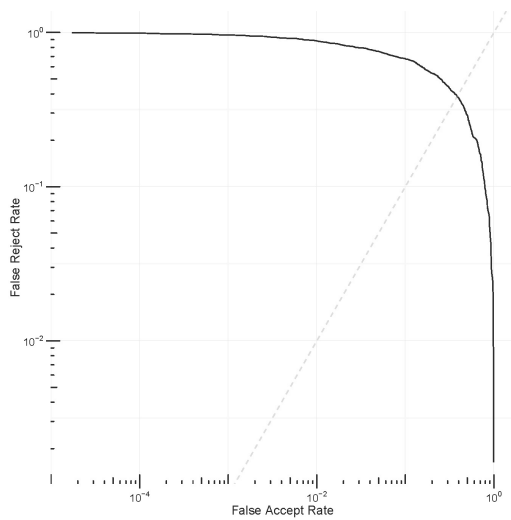
(b) S4F



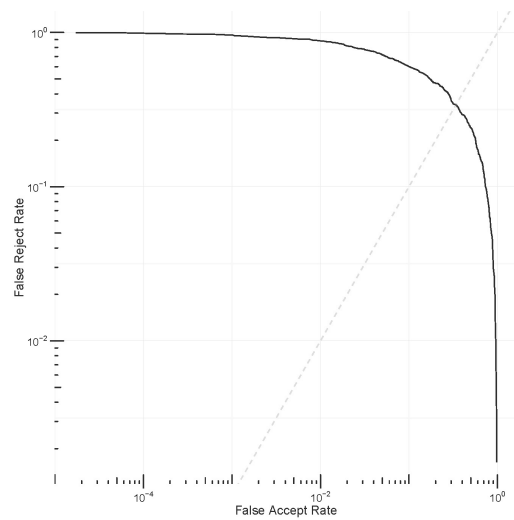
(c) CohortLDA



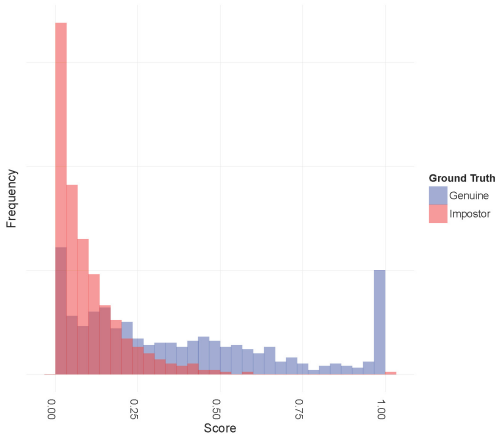
(d) LRPCA



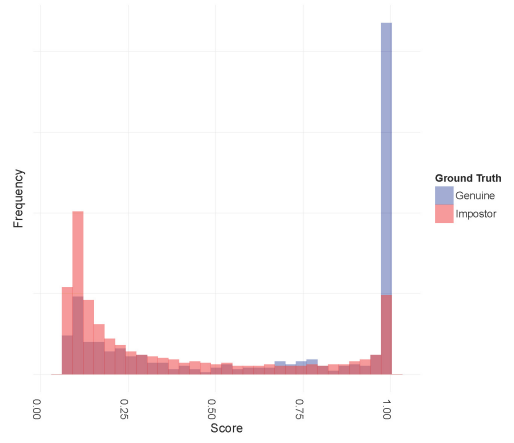
(e) LDA



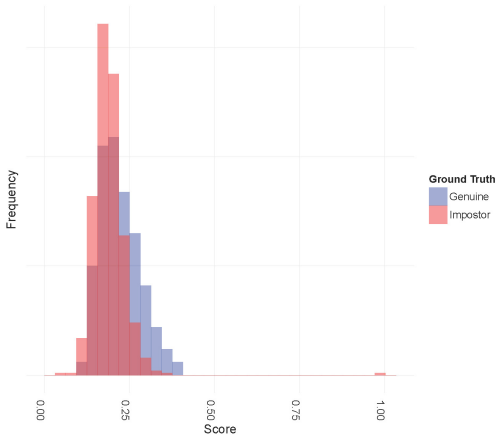
(f) PCA



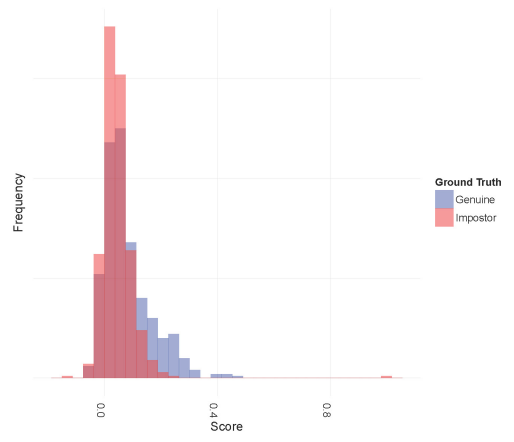
(a) Cognitec



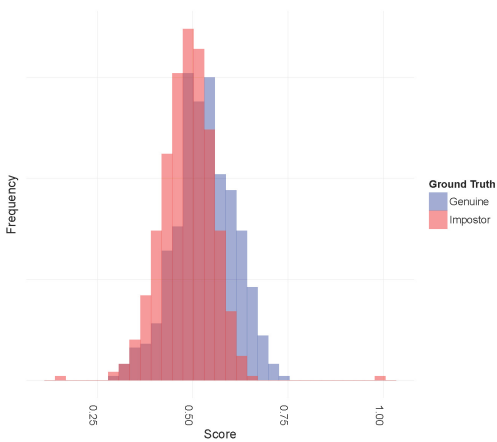
(b) S4F



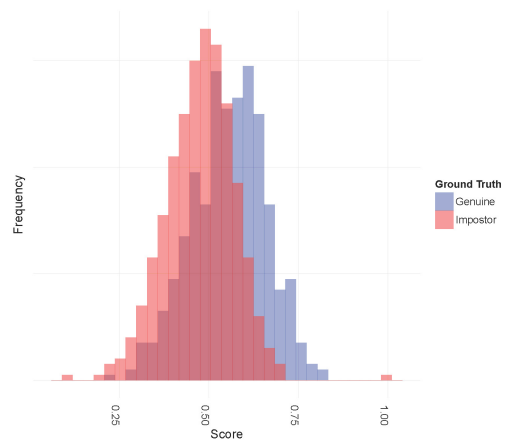
(c) CohortLDA



(d) LRPCA

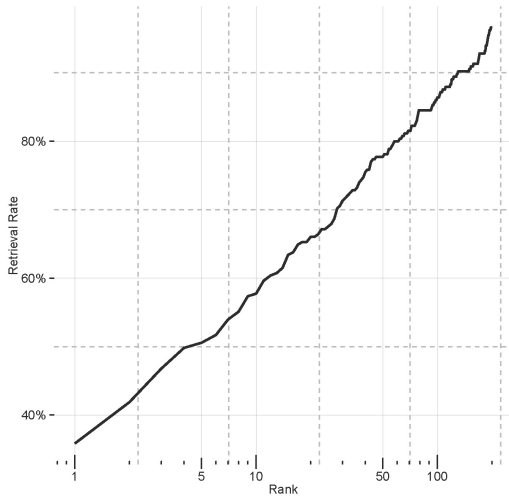


(e) LDA

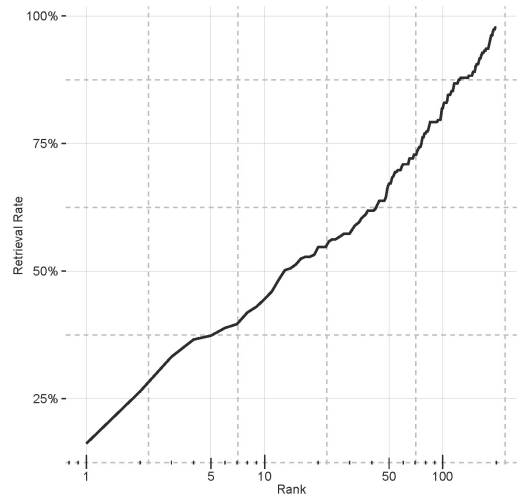


(f) PCA

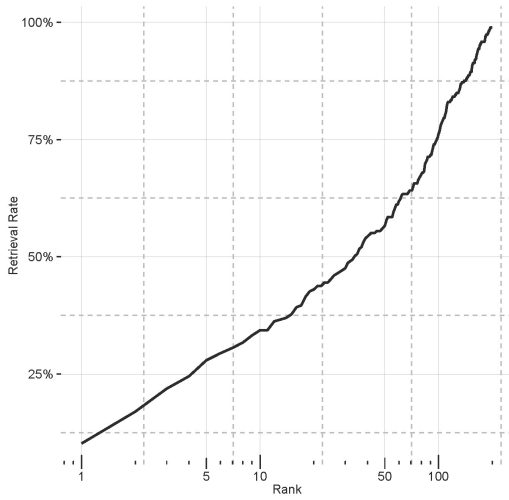
Figure 6.65: Experiment 4 9-11 to 12-14 - Score Histogram



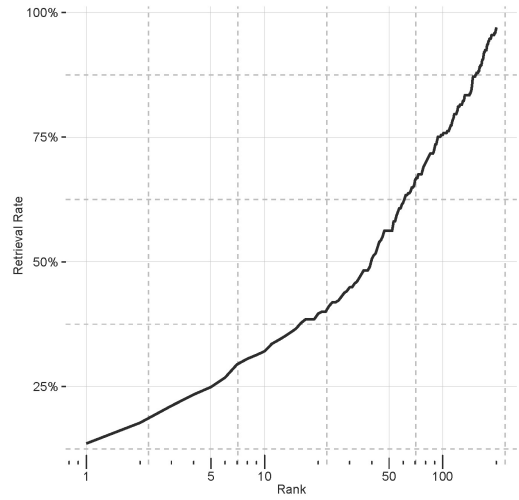
(a) Cognitec



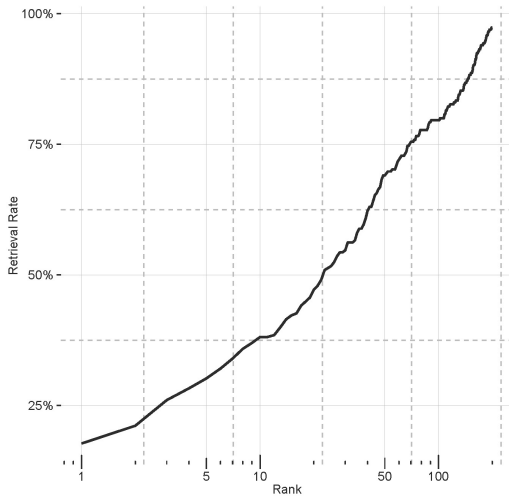
(b) S4F



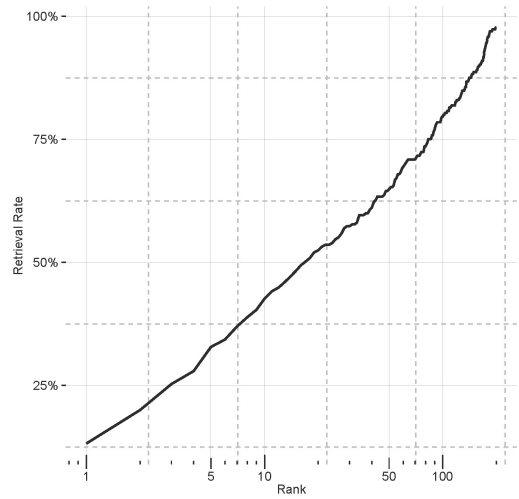
(c) CohortLDA



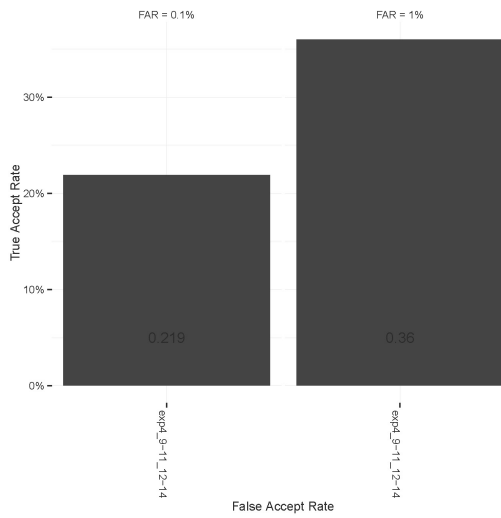
(d) LRPCA



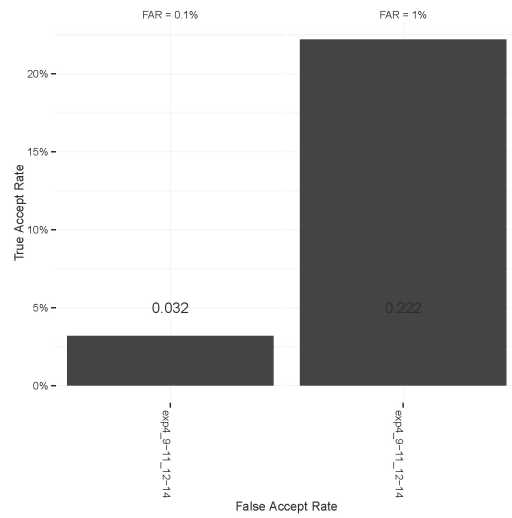
(e) LDA



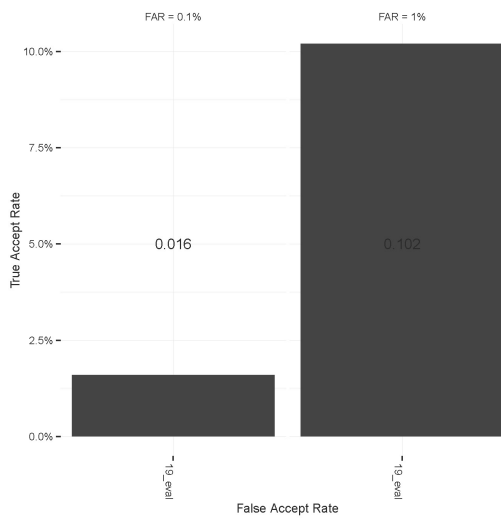
(f) PCA



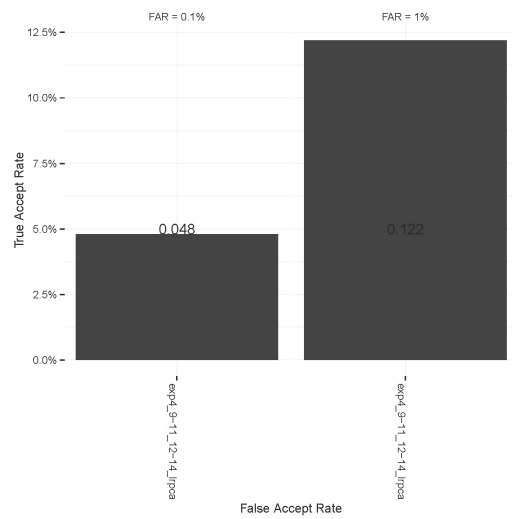
(a) Cognitec



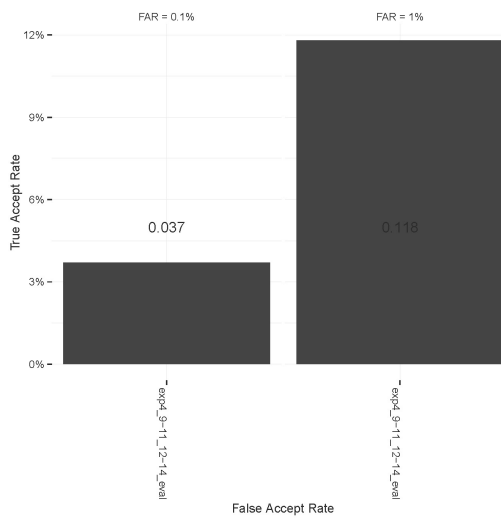
(b) S4F



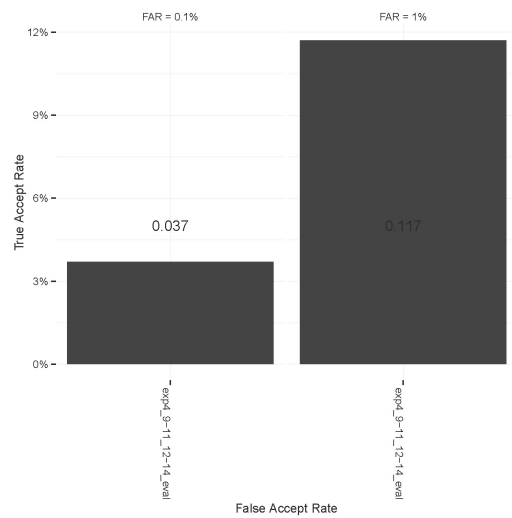
(c) CohortLDA



(d) LRPCA

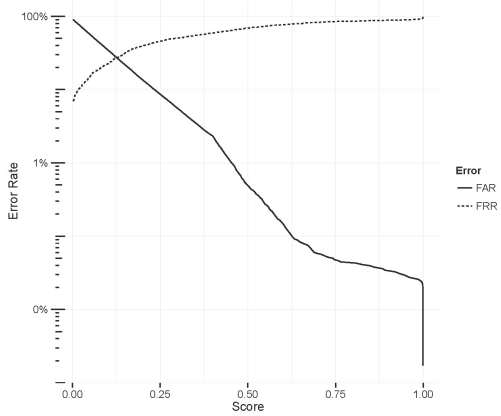


(e) LDA

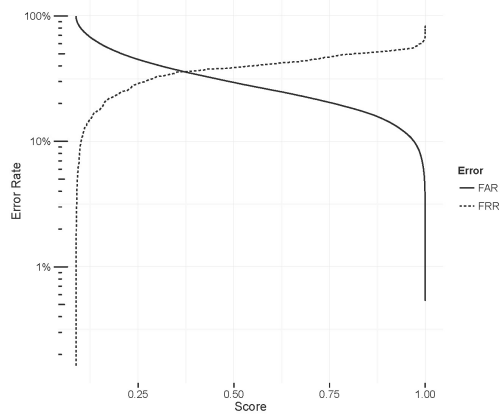


(f) PCA

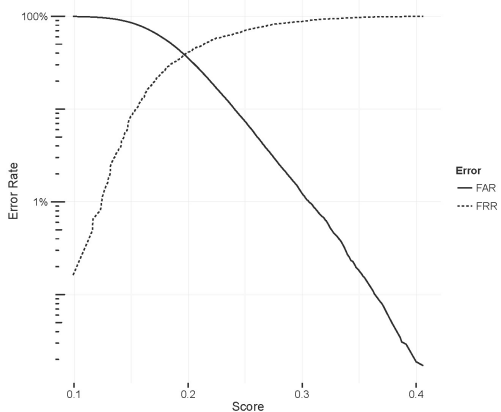
Figure 6.67: Experiment 4 9-11 to 12-14 - ROC Scores



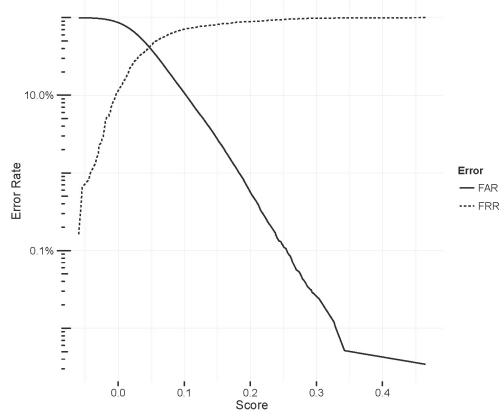
(a) Cognitec



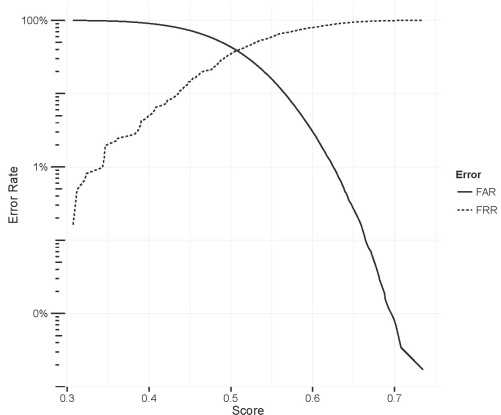
(b) S4F



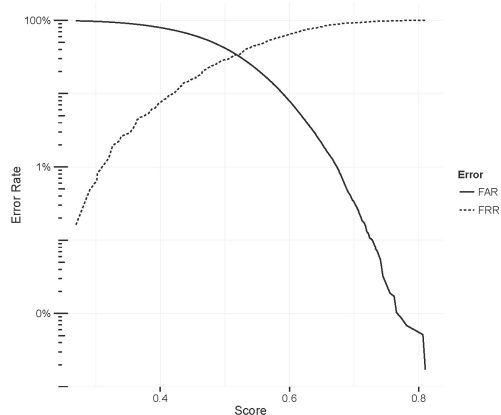
(c) CohortLDA



(d) LRPCA



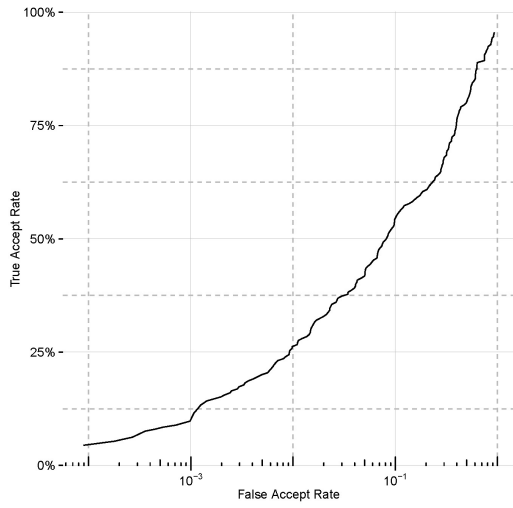
(e) LDA



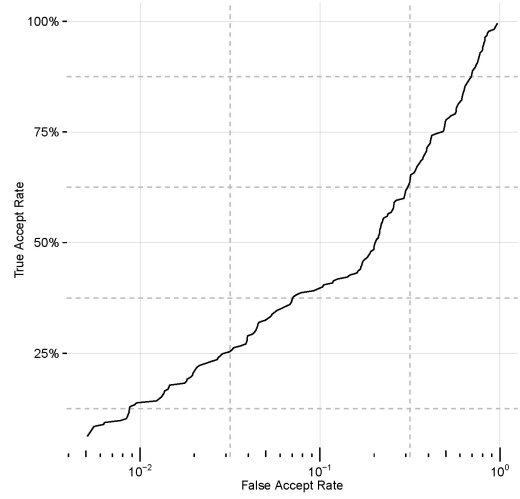
(f) PCA

APPENDICES J

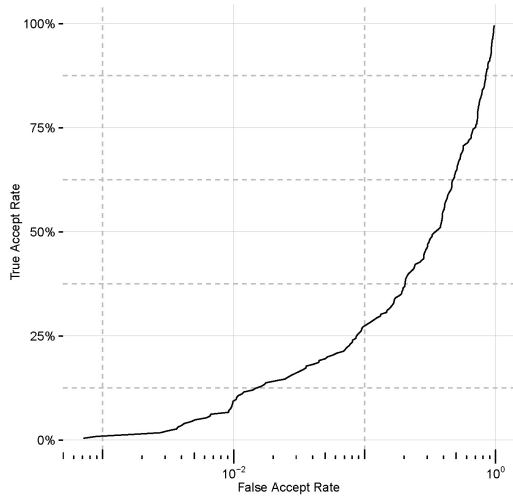
Experiment 4 - 9-11 to 15-17 Results



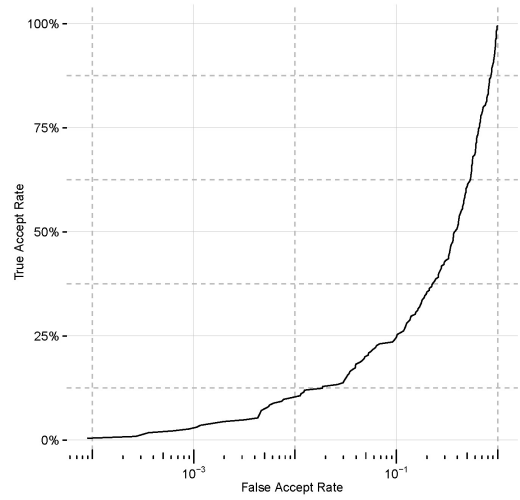
(a) Cognittec



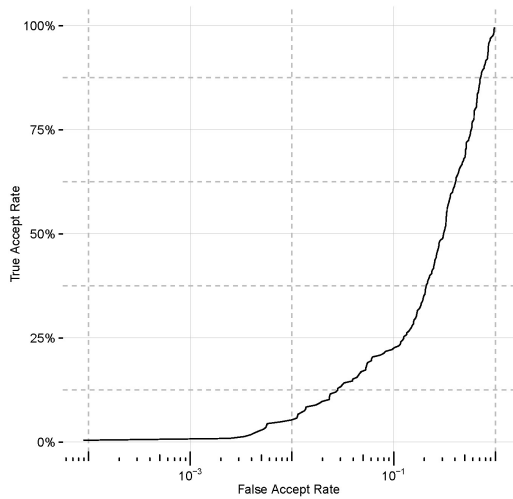
(b) S4F



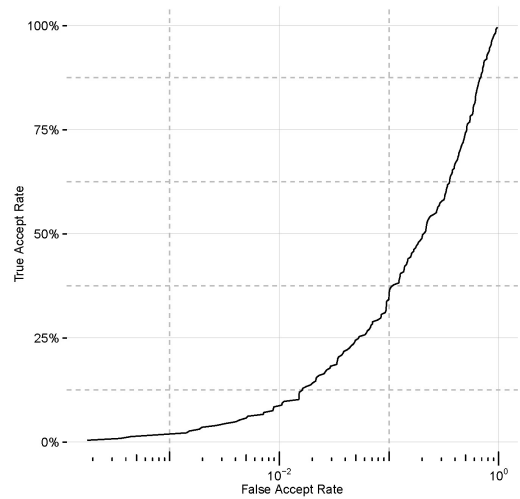
(c) CohortLDA



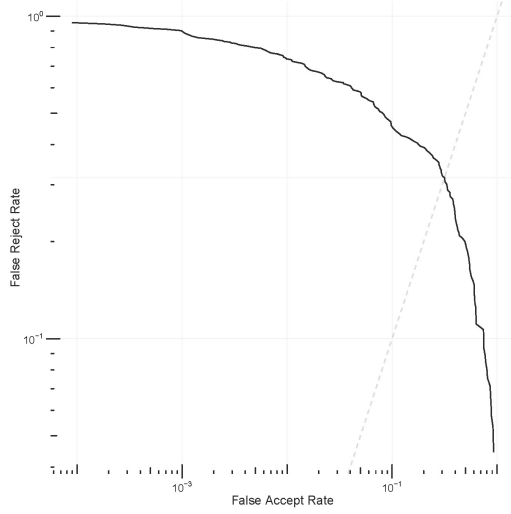
(d) LRPCA



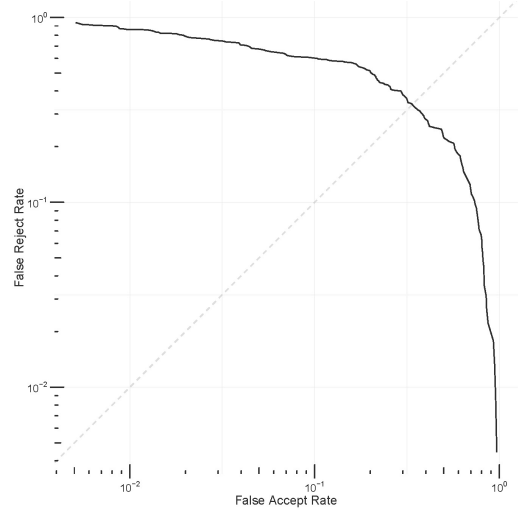
(e) LDA



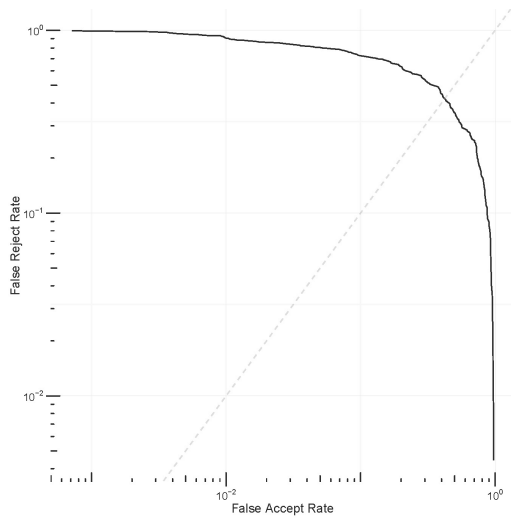
(f) PCA



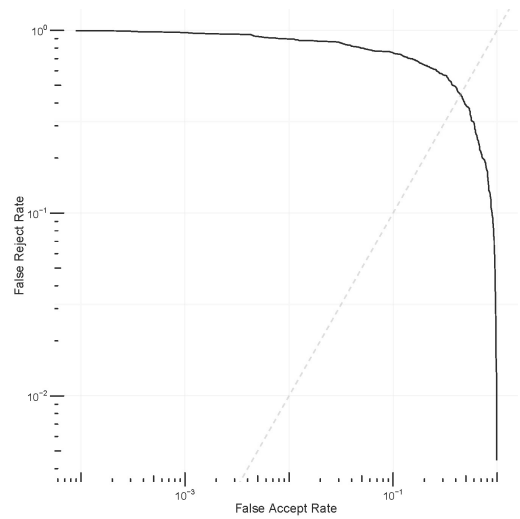
(a) Cognittec



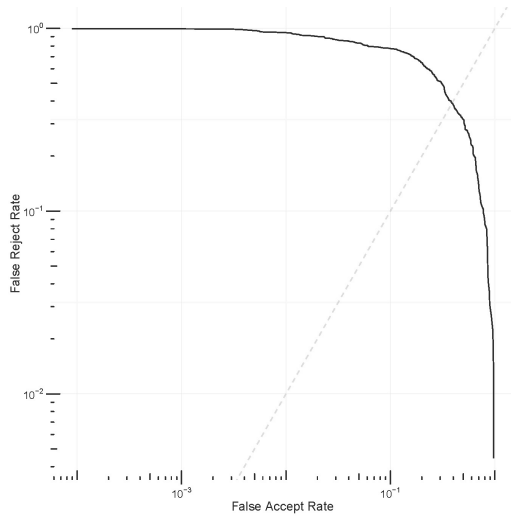
(b) S4F



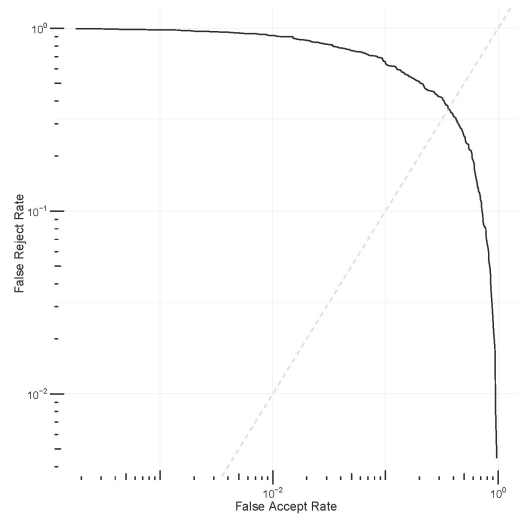
(c) CohortLDA



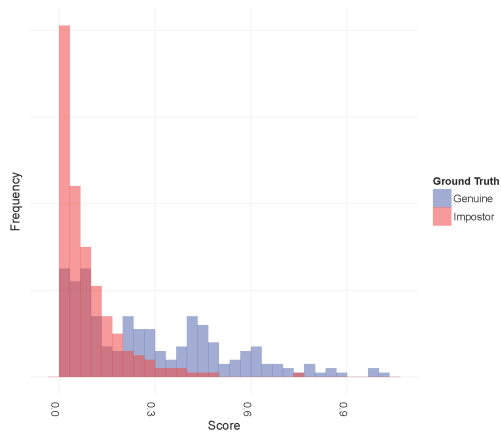
(d) LRPCA



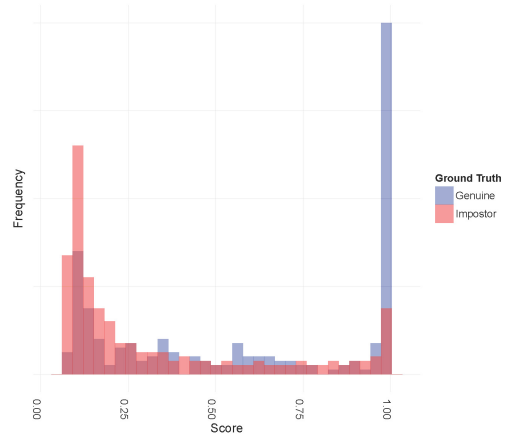
(e) LDA



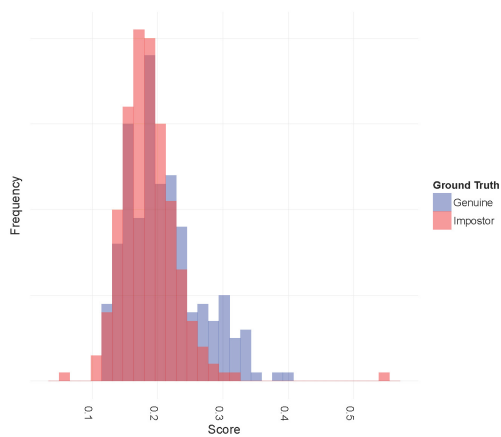
(f) PCA



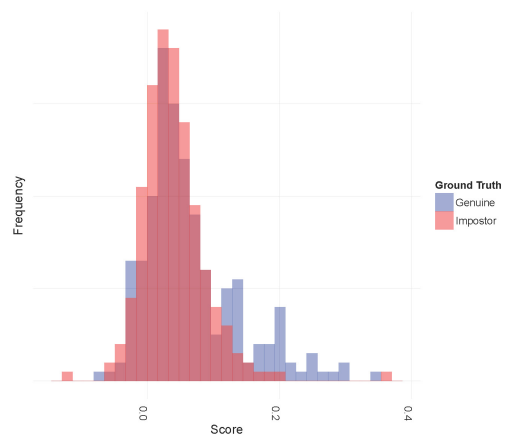
(a) Cognitec



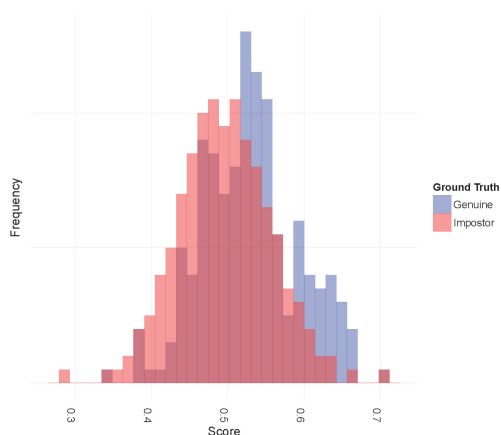
(b) S4F



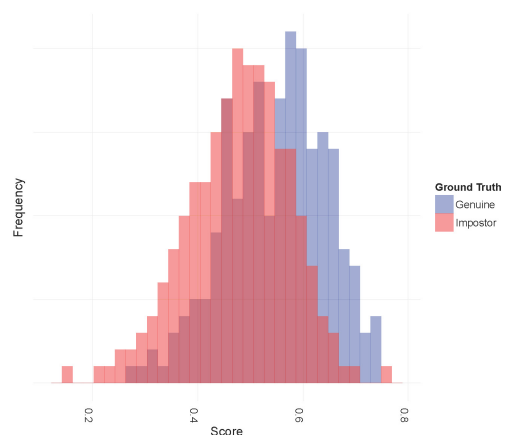
(c) CohortLDA



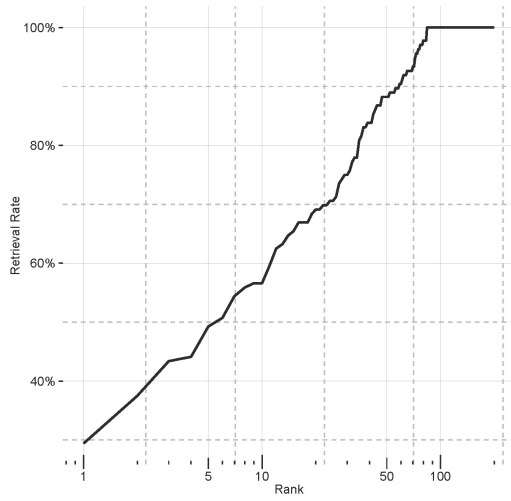
(d) LRPCA



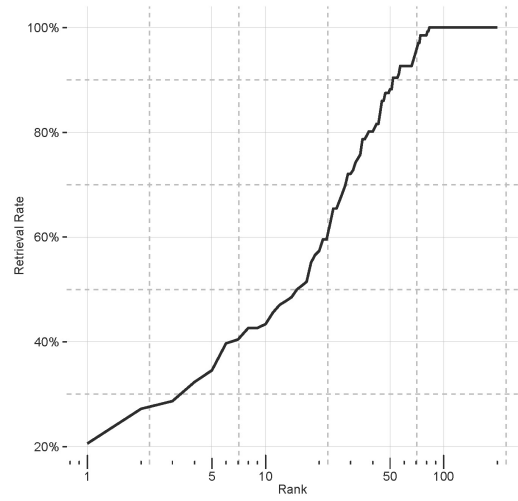
(e) LDA



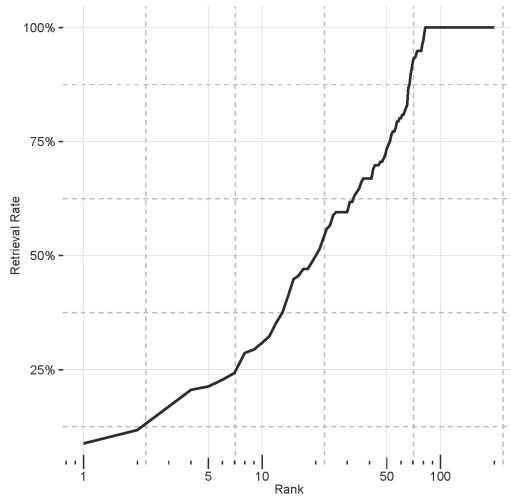
(f) PCA



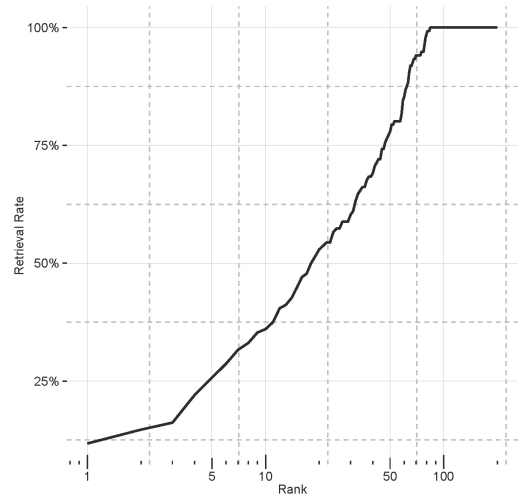
(a) Cognitec



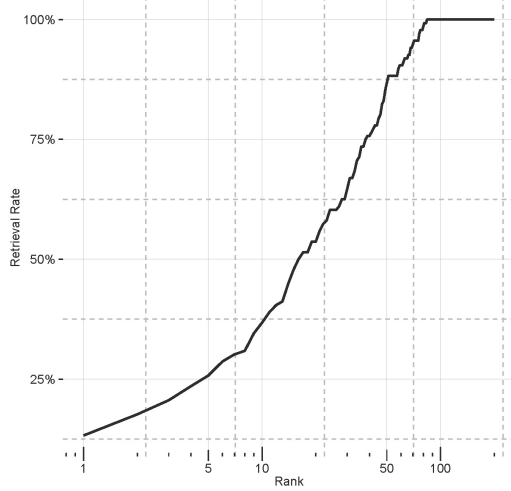
(b) S4F



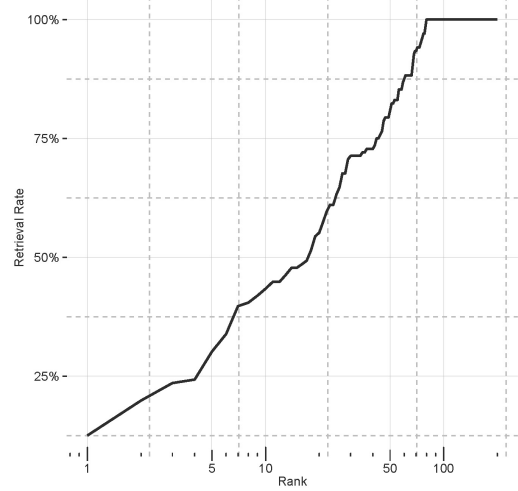
(c) CohortLDA



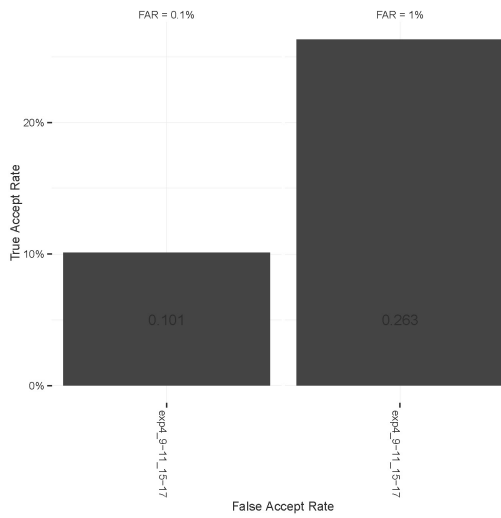
(d) LRPCA



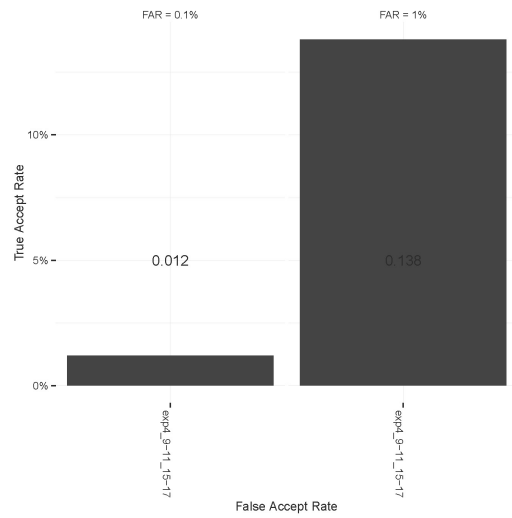
(e) LDA



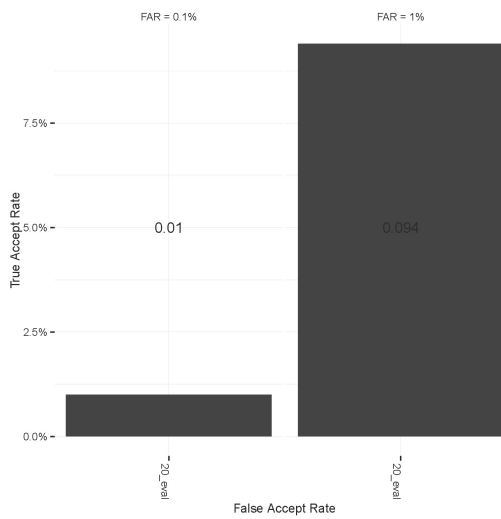
(f) PCA



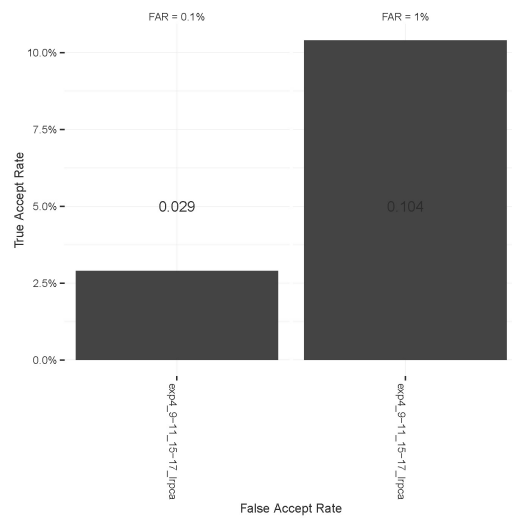
(a) Cognitec



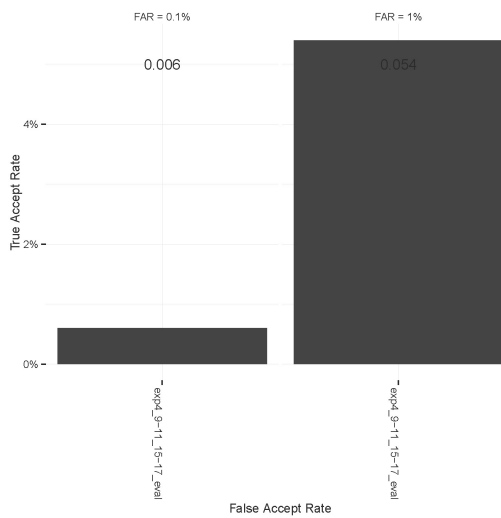
(b) S4F



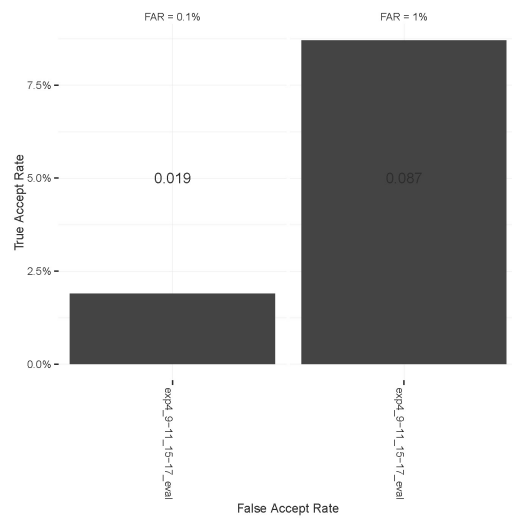
(c) CohortLDA



(d) LRPCA

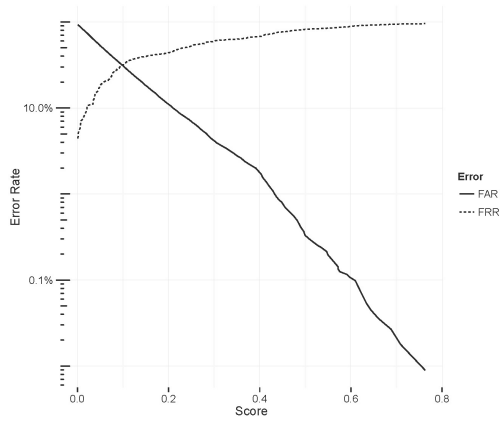


(e) LDA

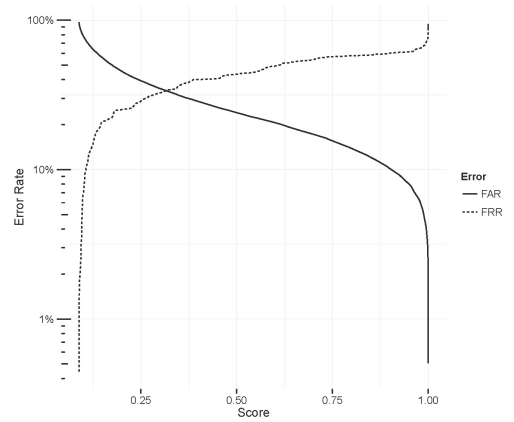


(f) PCA

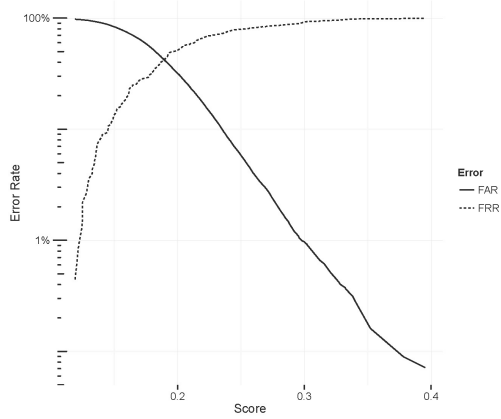
Figure 6.73: Experiment 4 9-11 to 15-17 - ROC Scores



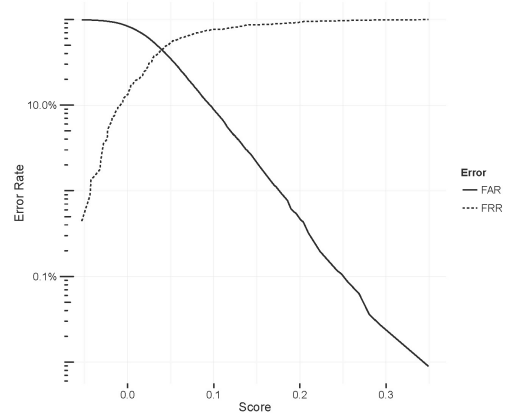
(a) Cognitec



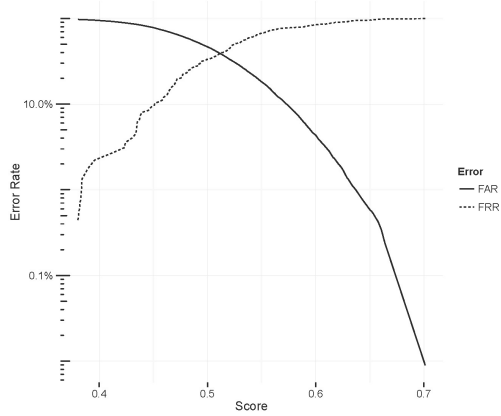
(b) S4F



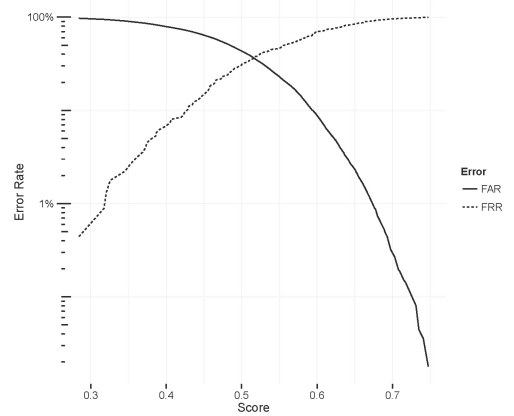
(c) CohortLDA



(d) LRPCA



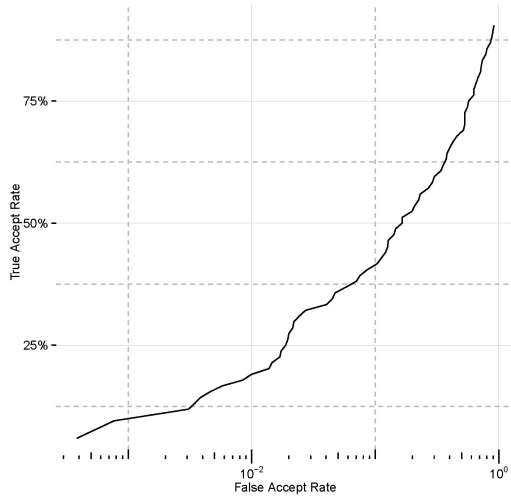
(e) LDA



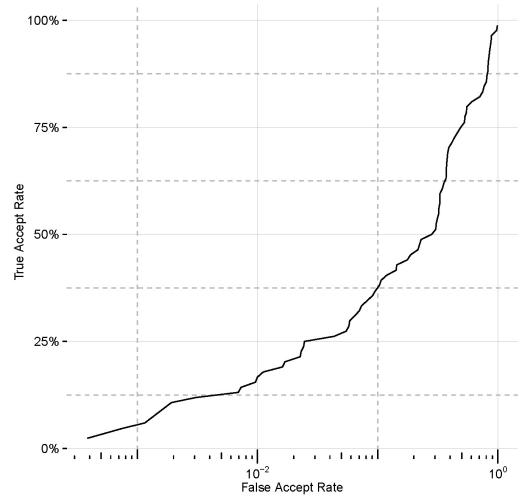
(f) PCA

APPENDICES K

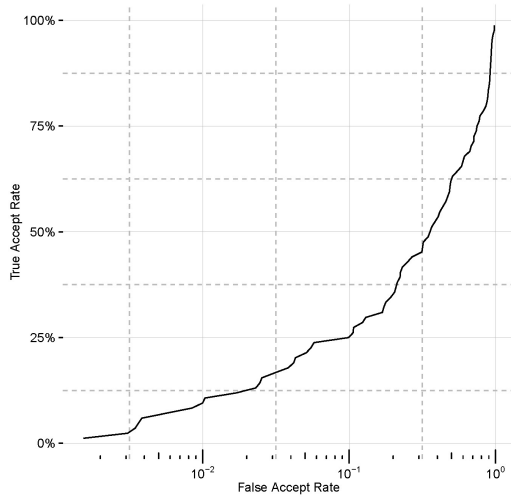
Experiment 4 - 9-11 to 18+ Results



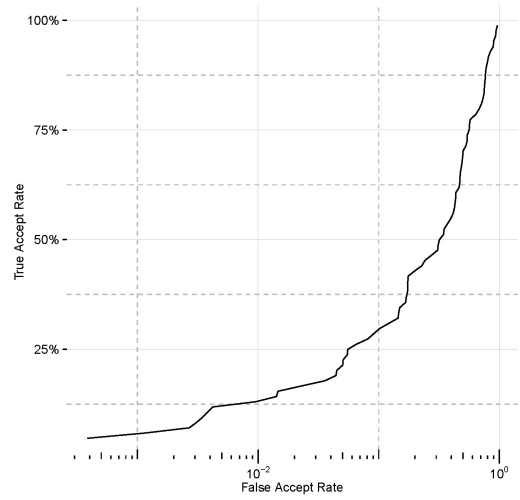
(a) Cognittec



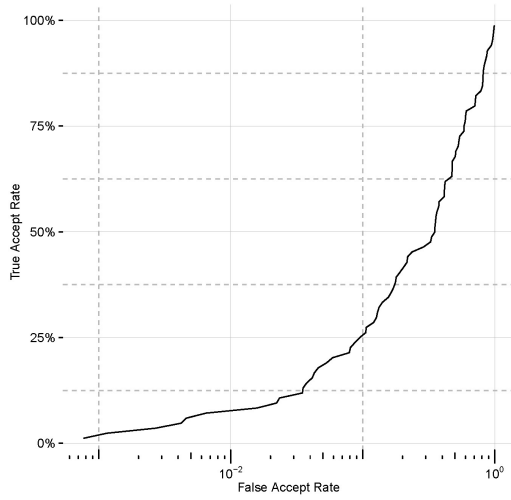
(b) S4F



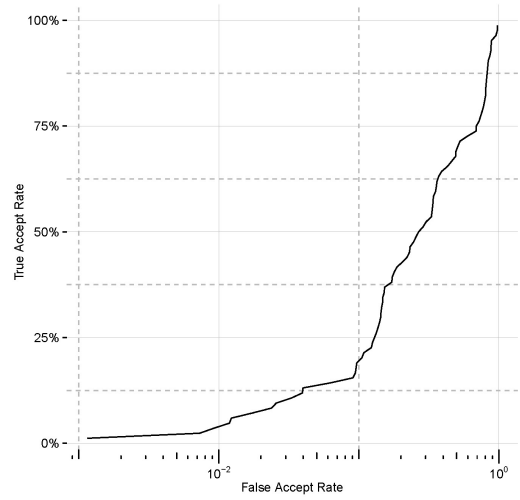
(c) CohortLDA



(d) LRPCA

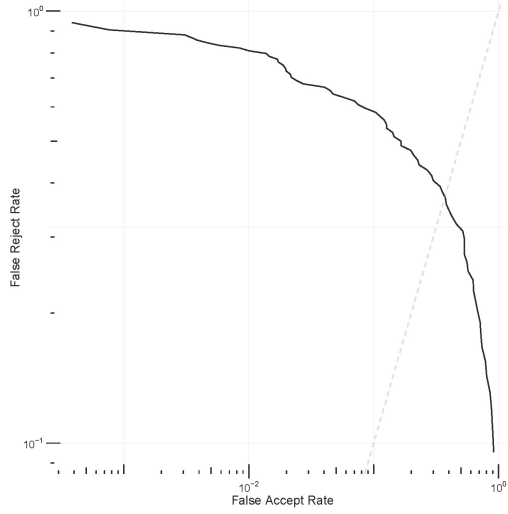


(e) LDA

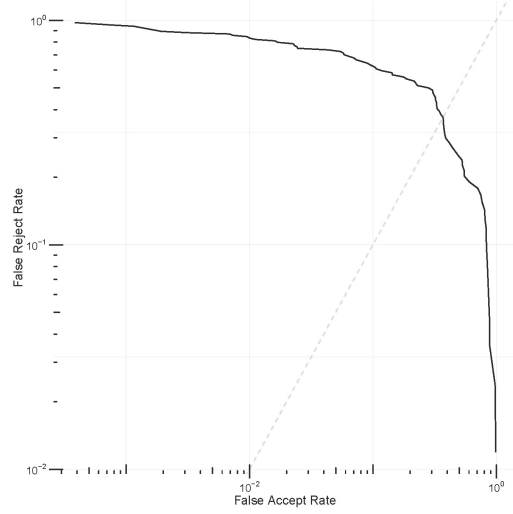


(f) PCA

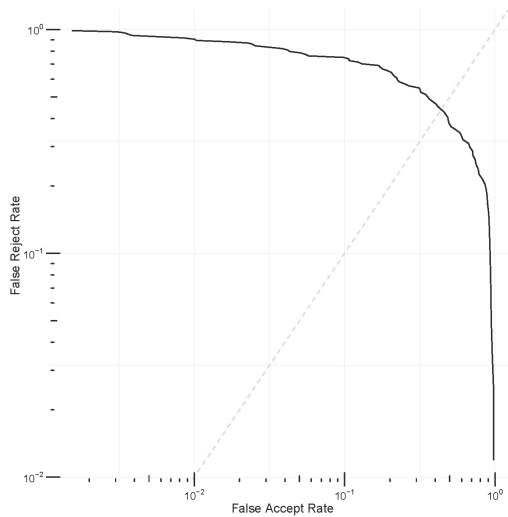
Figure 6.75: Experiment 4 9-11 to 18+ - ROC



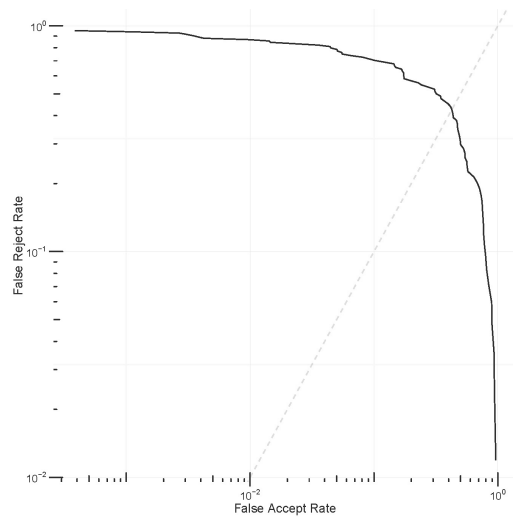
(a) Cognitec



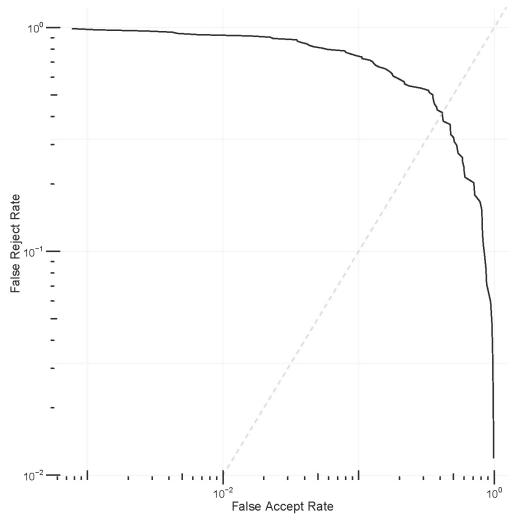
(b) S4F



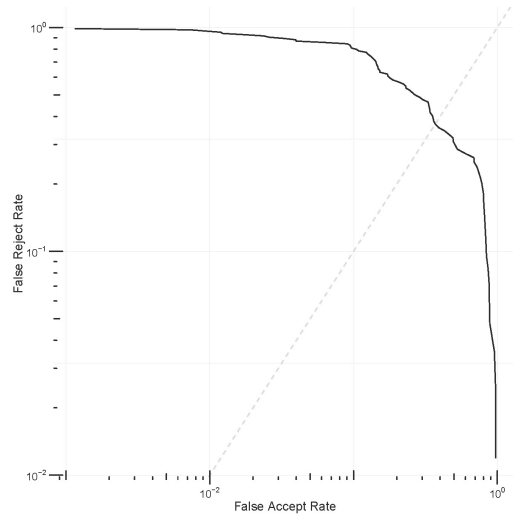
(c) CohortLDA



(d) LRPCA

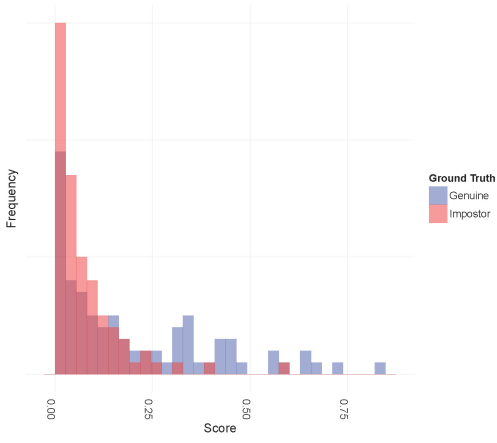


(e) LDA

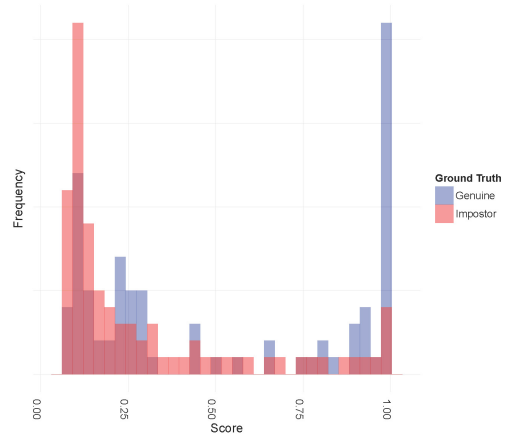


(f) PCA

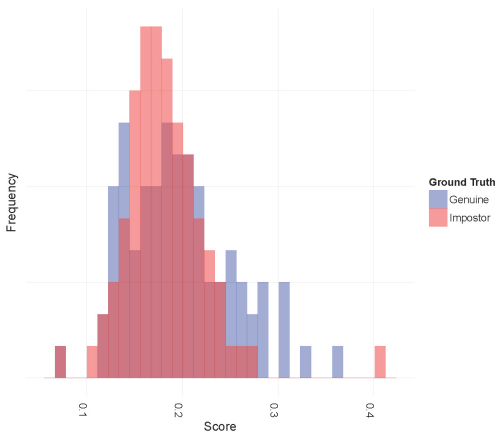
Figure 6.76: Experiment 4 9-11 to 18+ - DET



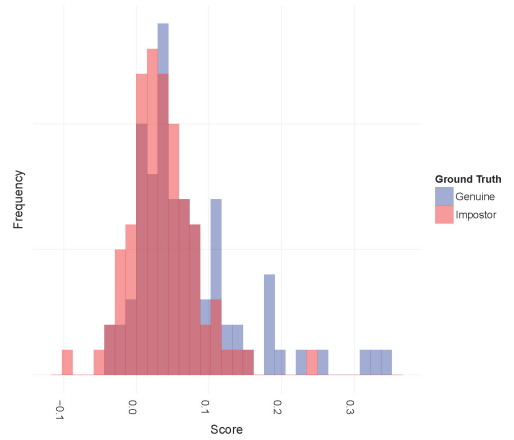
(a) Cognitec



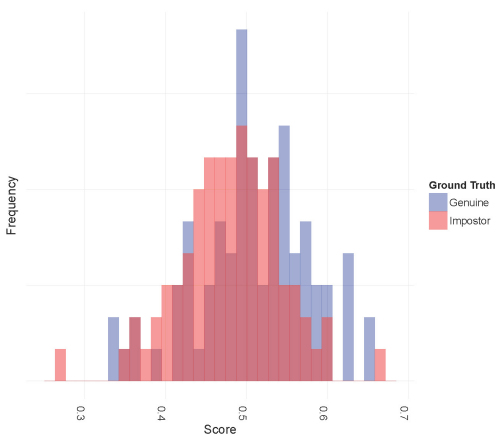
(b) S4F



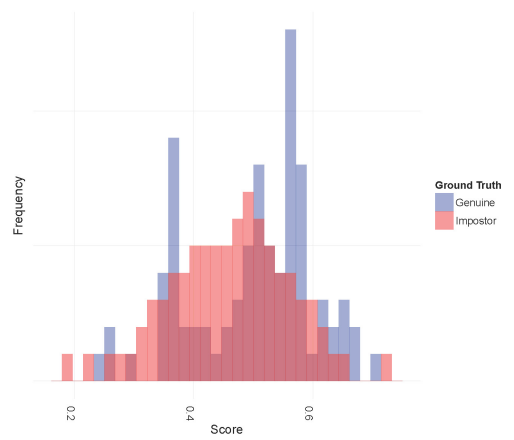
(c) CohortLDA



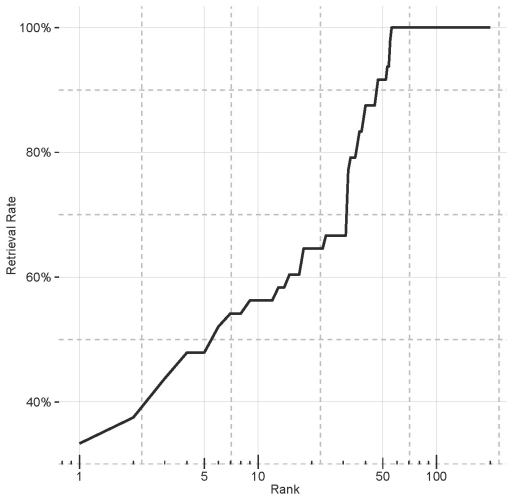
(d) LRPCA



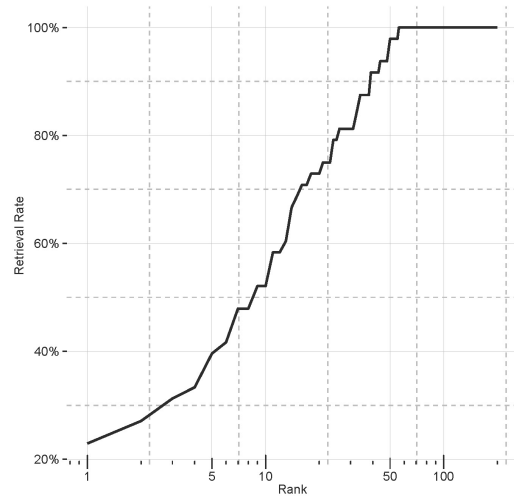
(e) LDA



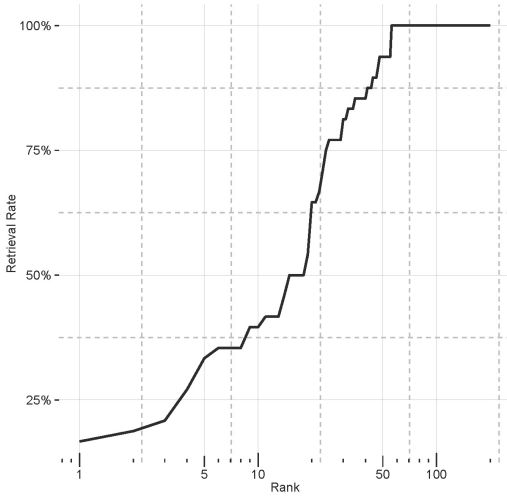
(f) PCA



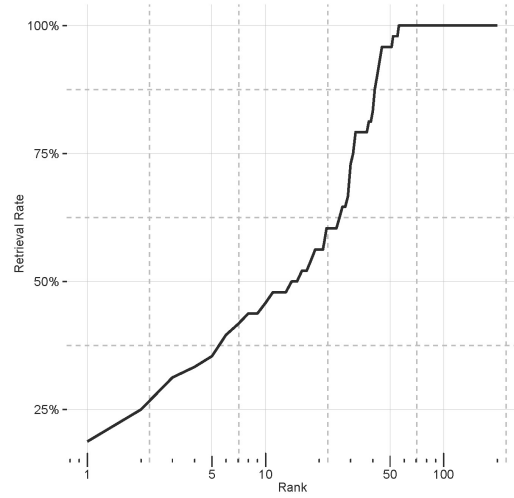
(a) Cognitec



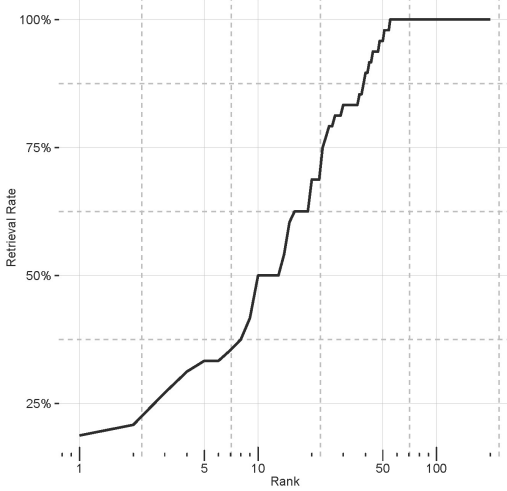
(b) S4F



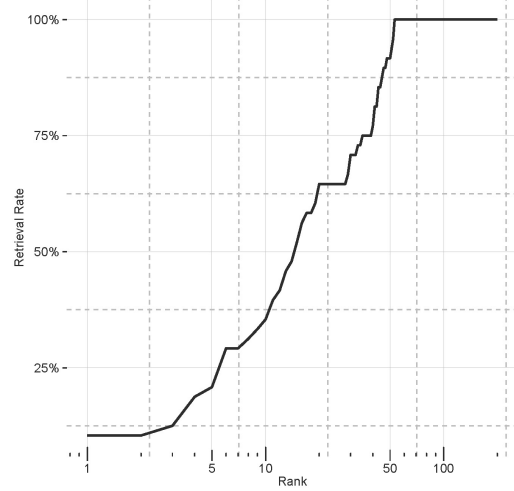
(c) CohortLDA



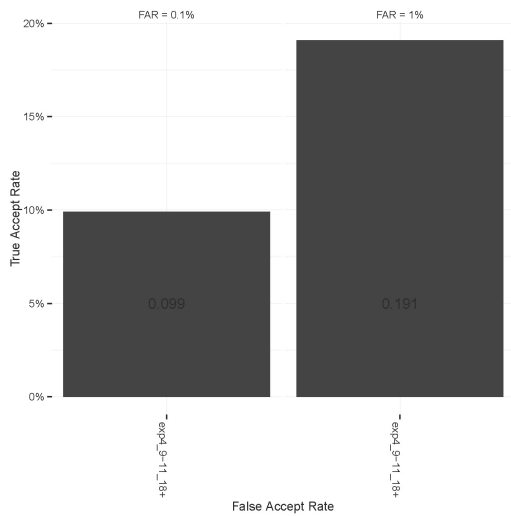
(d) LRPCA



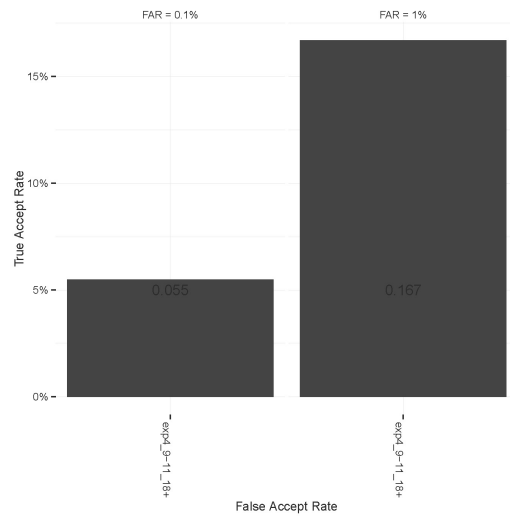
(e) LDA



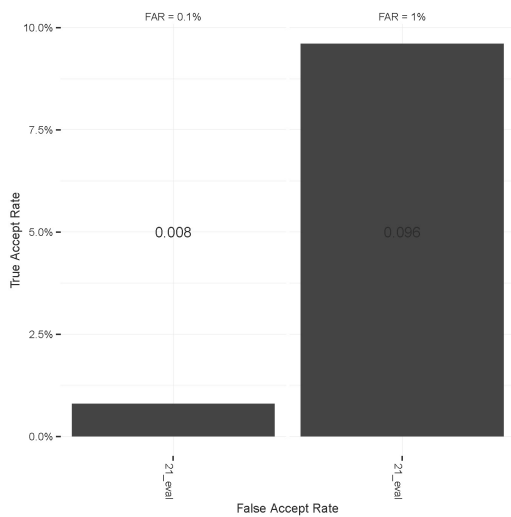
(f) PCA



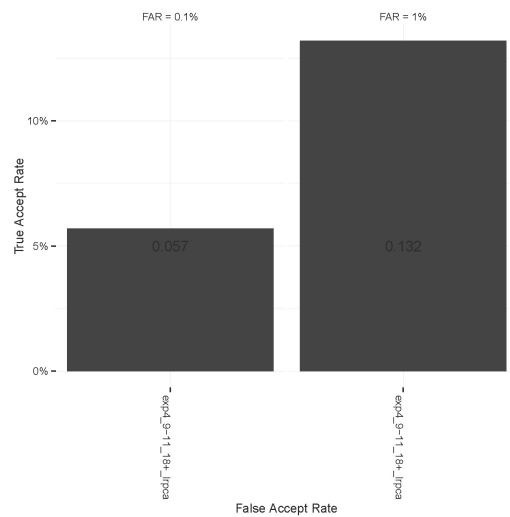
(a) Cognitex



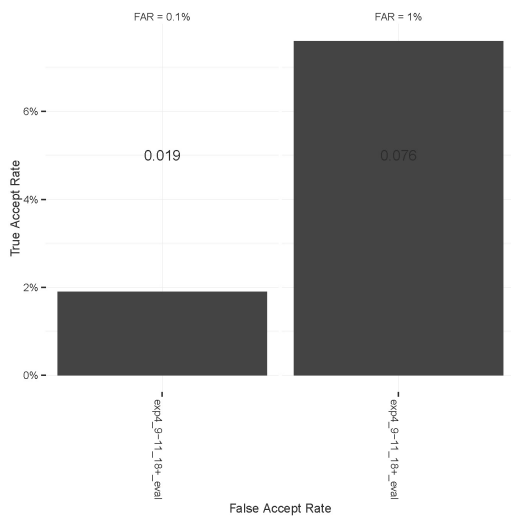
(b) S4F



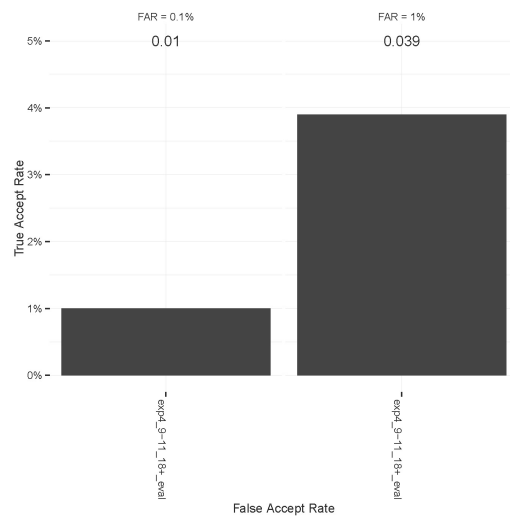
(c) CohortLDA



(d) LRPCA

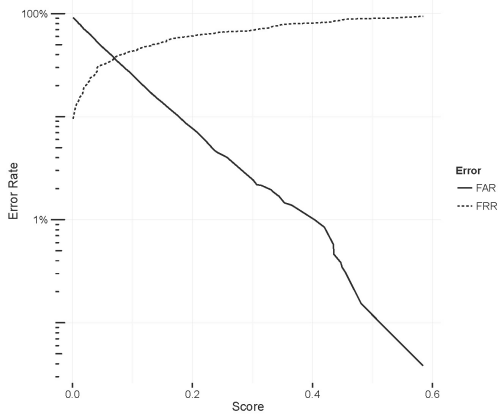


(e) LDA

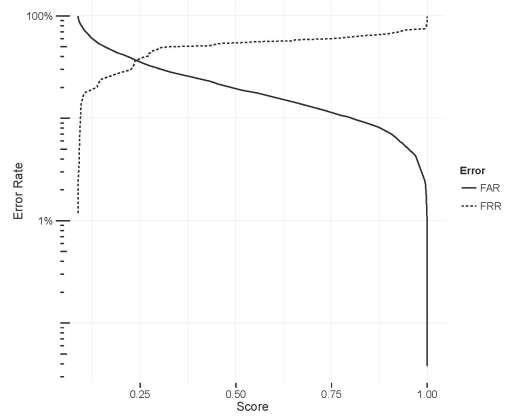


(f) PCA

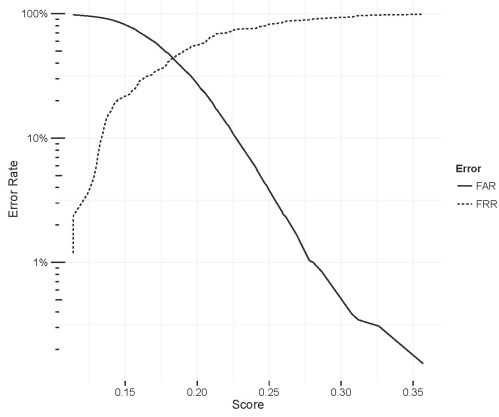
Figure 6.79: Experiment 4 9-11 to 18+ - ROC Scores



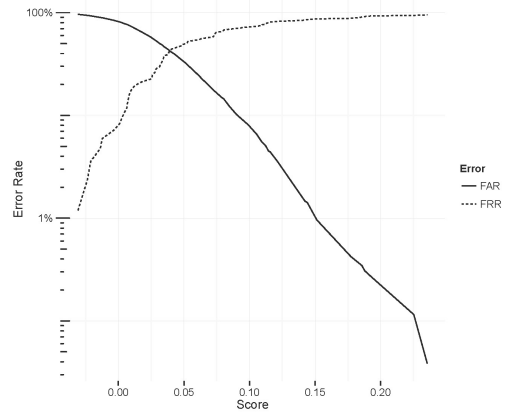
(a) Cognitec



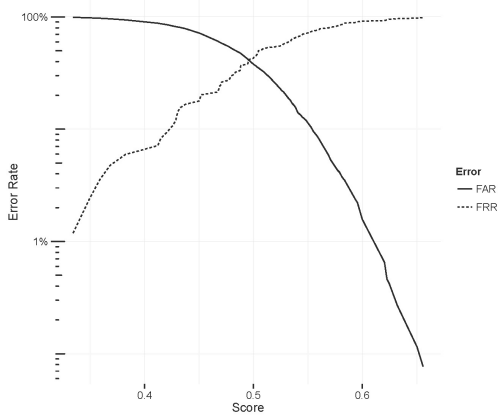
(b) S4F



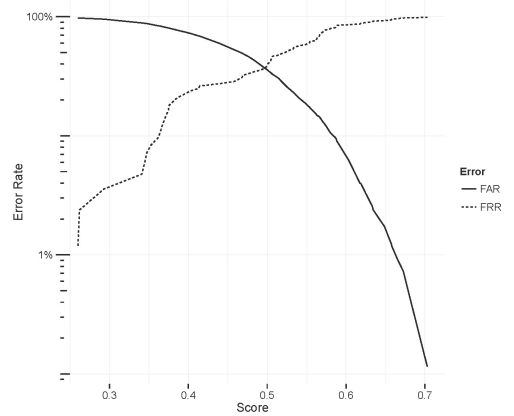
(c) CohortLDA



(d) LRPCA



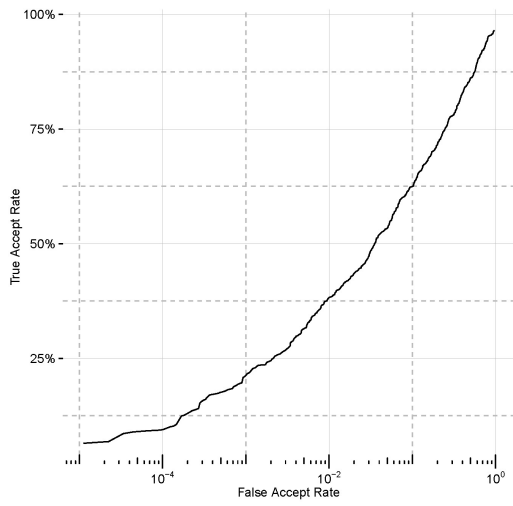
(e) LDA



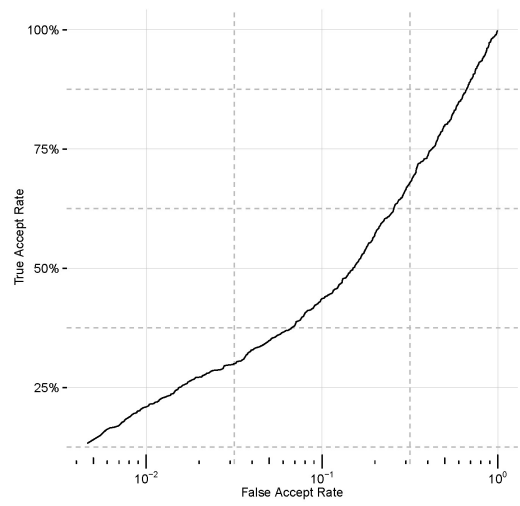
(f) PCA

APPENDICES L

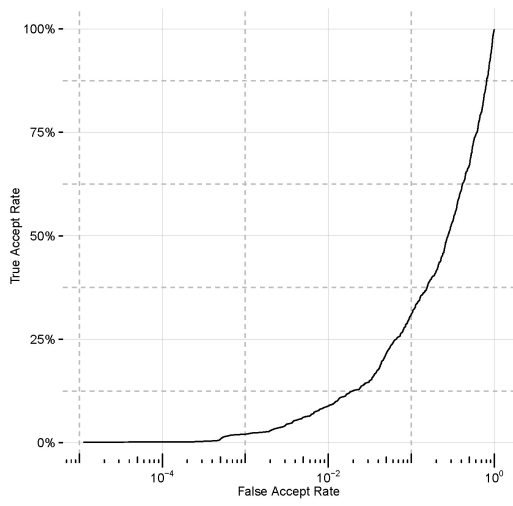
Experiment 4 - 12-14 to 15-17 Results



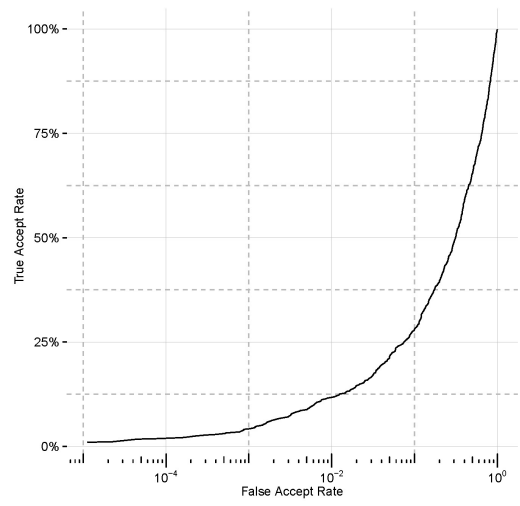
(a) Cognitec



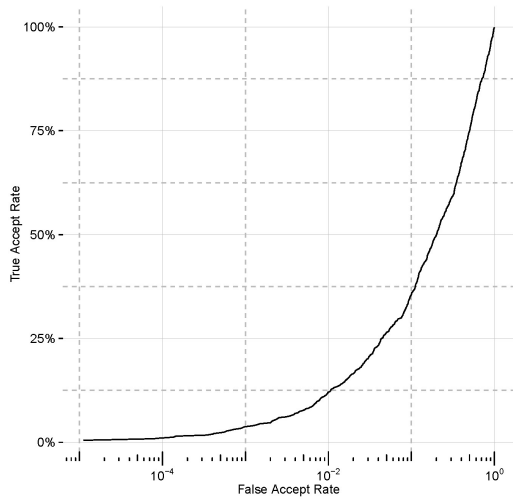
(b) S4F



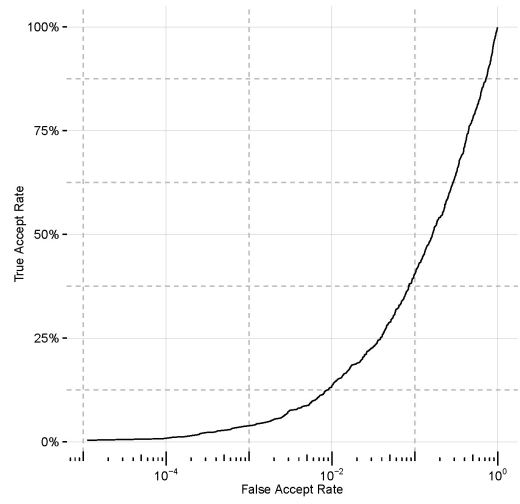
(c) CohortLDA



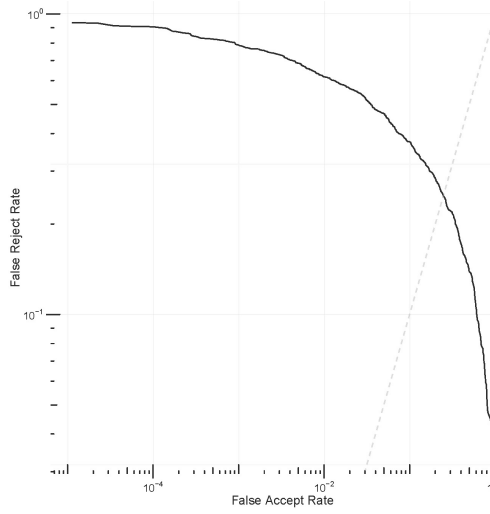
(d) LRPCA



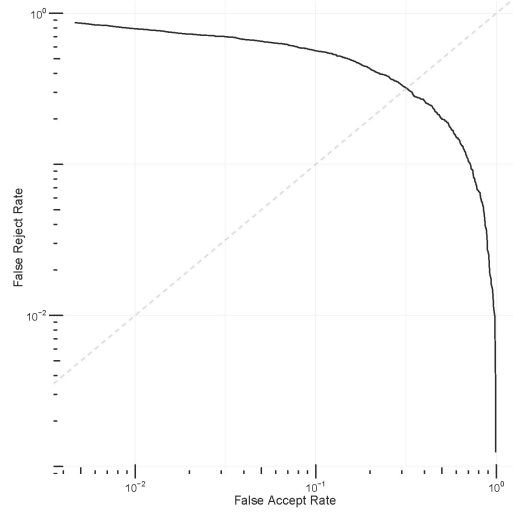
(e) LDA



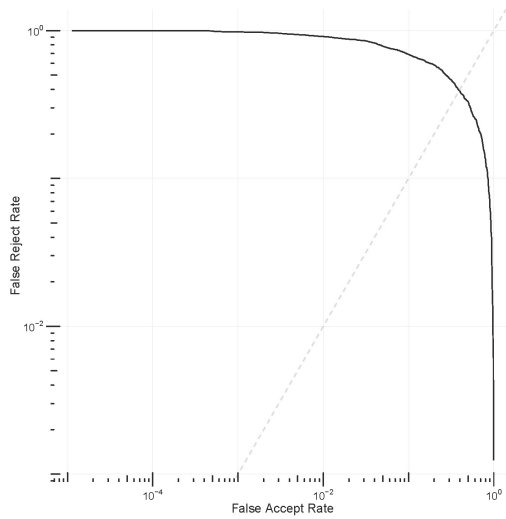
(f) PCA



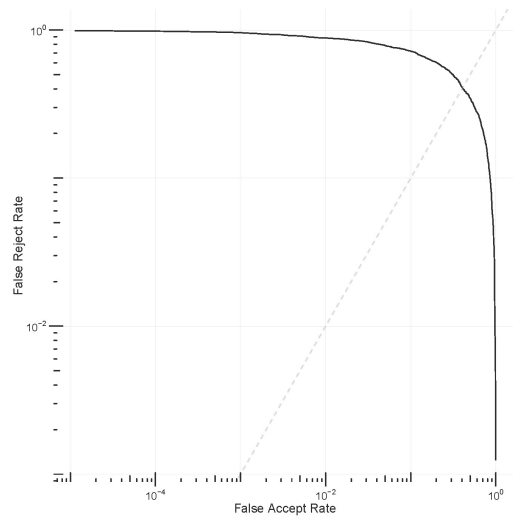
(a) Cognitec



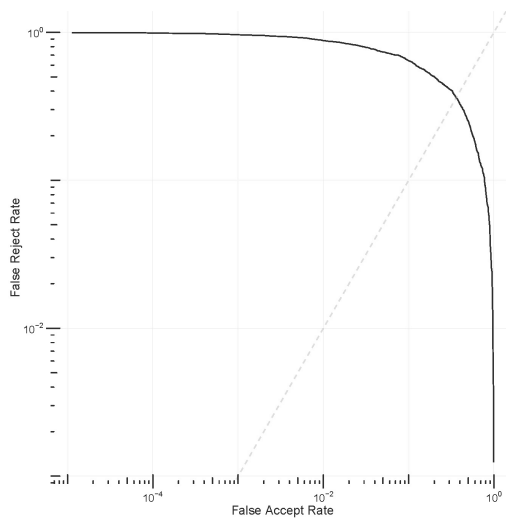
(b) S4F



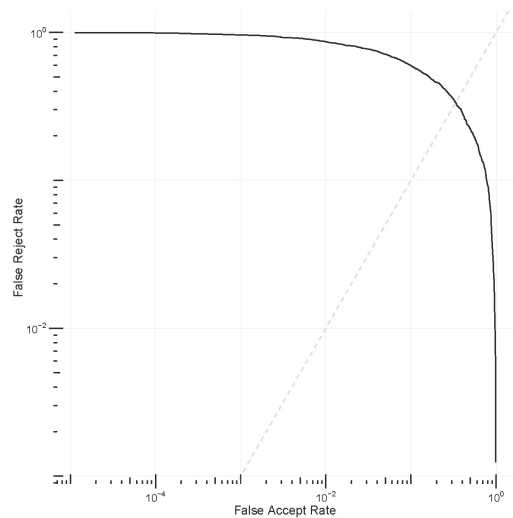
(c) CohortLDA



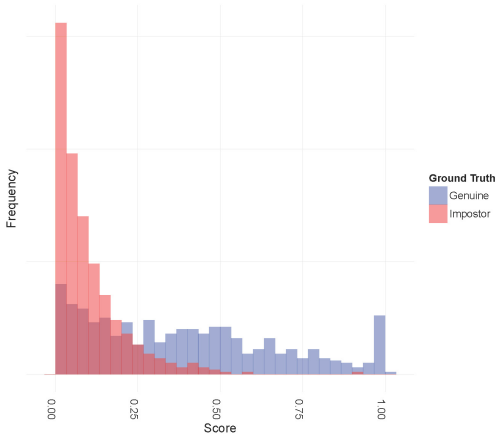
(d) LRPCA



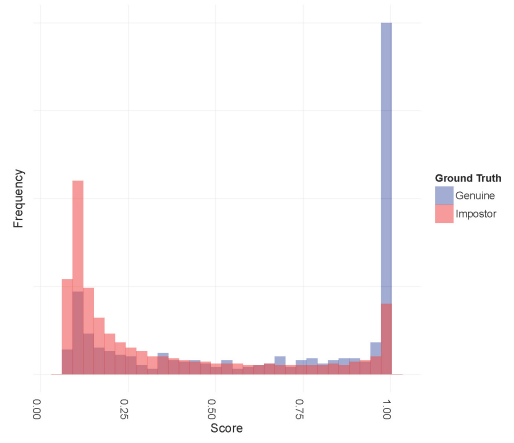
(e) LDA



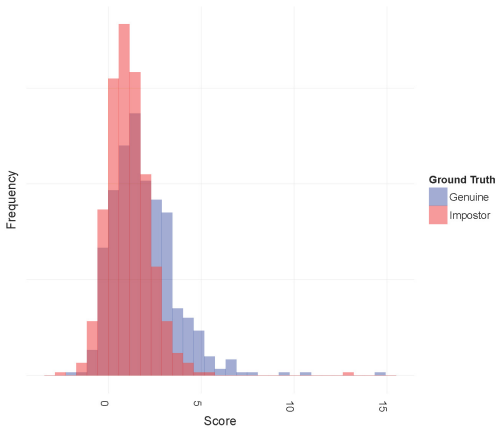
(f) PCA



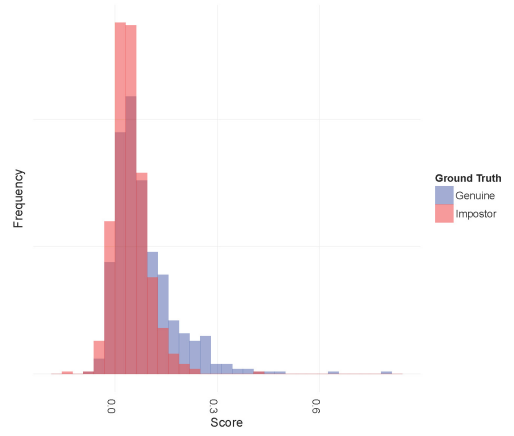
(a) Cognitec



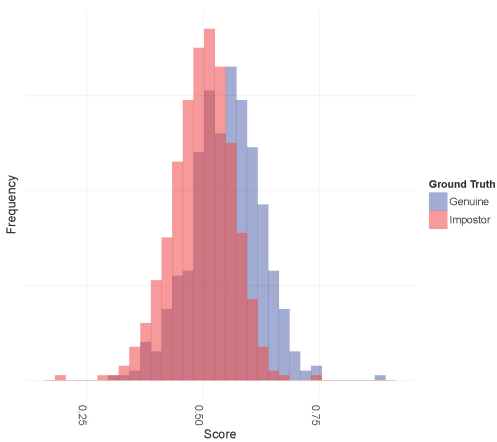
(b) S4F



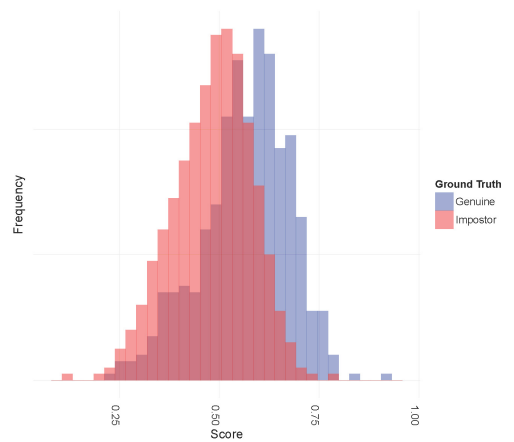
(c) CohortLDA



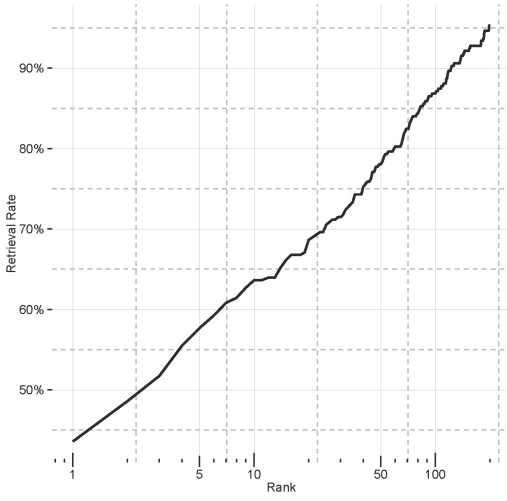
(d) LRPCA



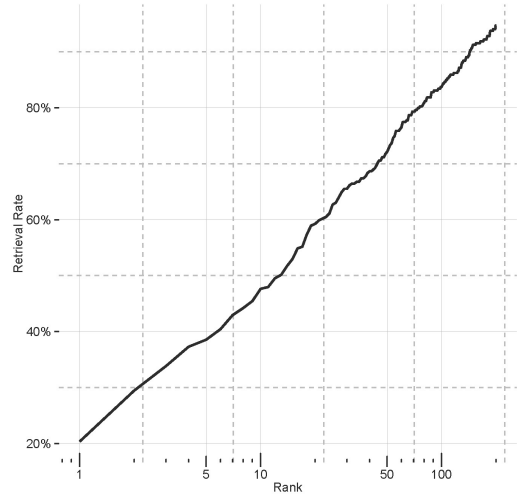
(e) LDA



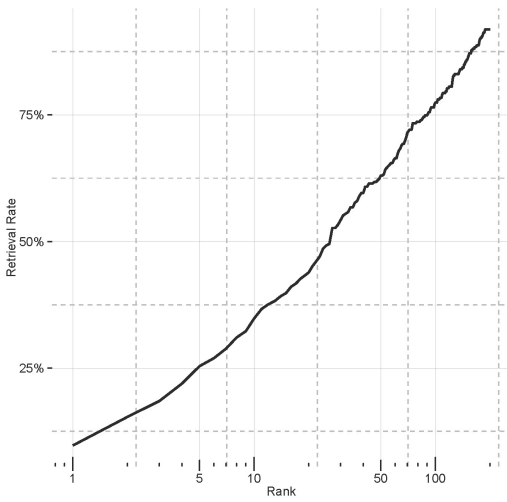
(f) PCA



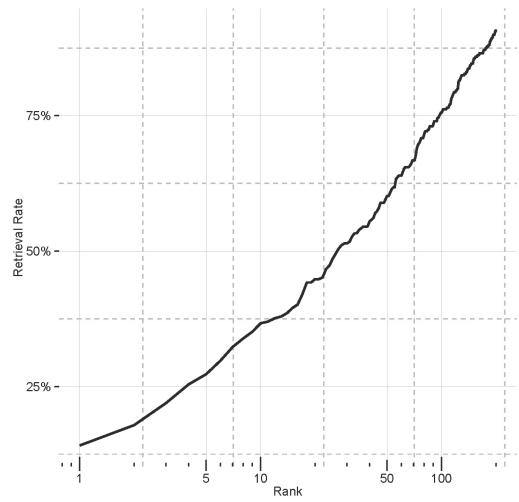
(a) Cognitec



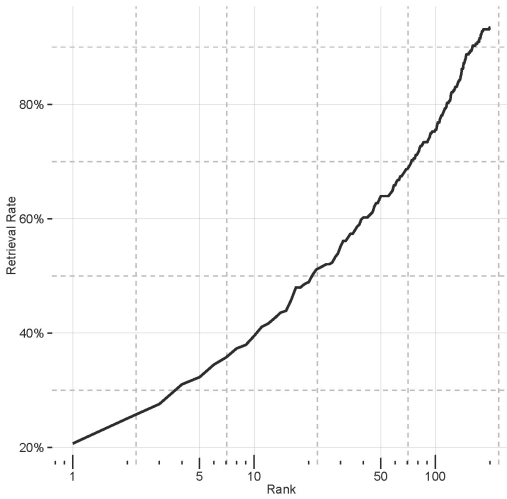
(b) S4F



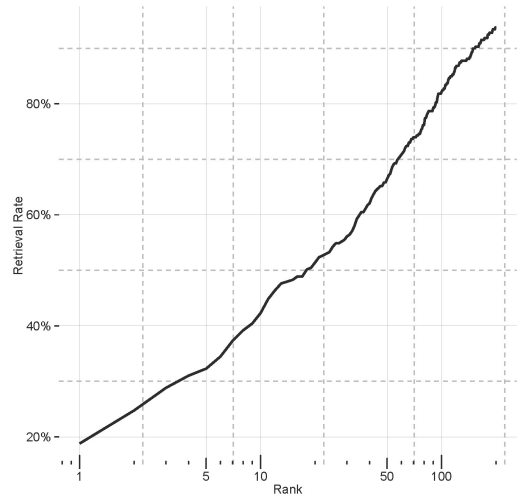
(c) CohortLDA



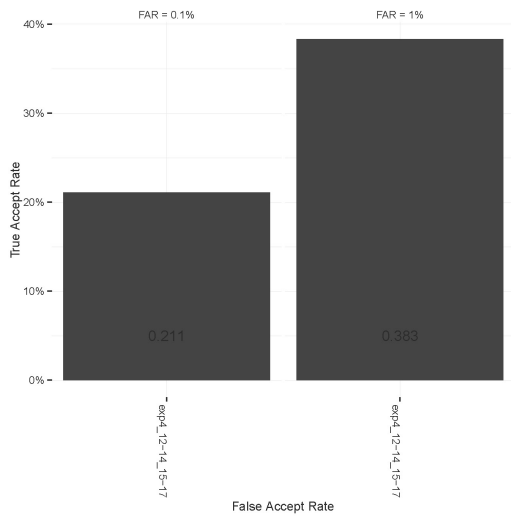
(d) LRPCA



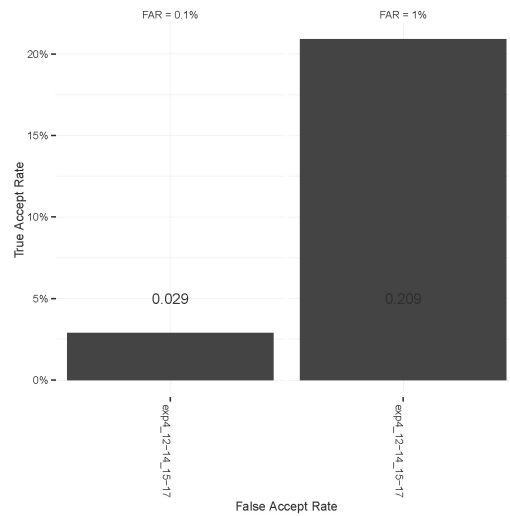
(e) LDA



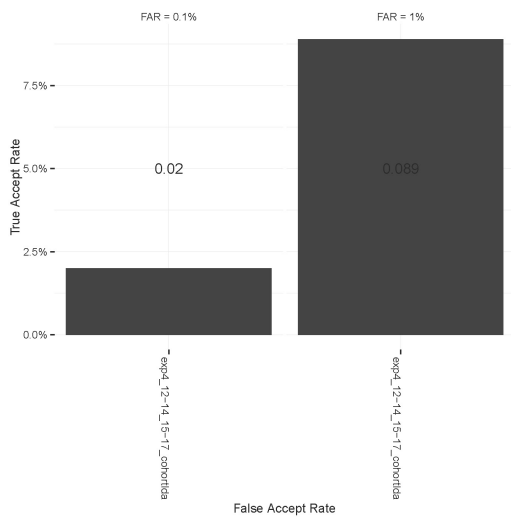
(f) PCA



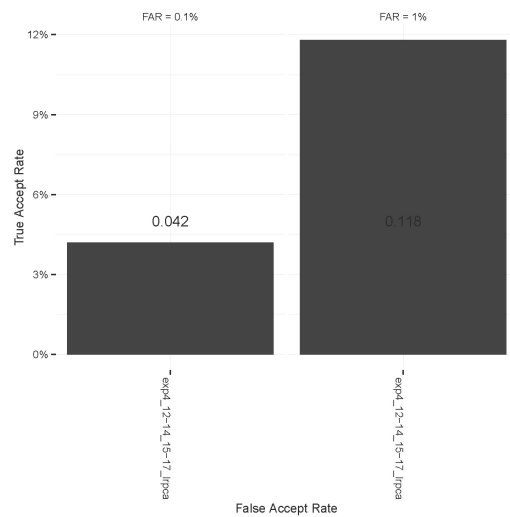
(a) Cognitec



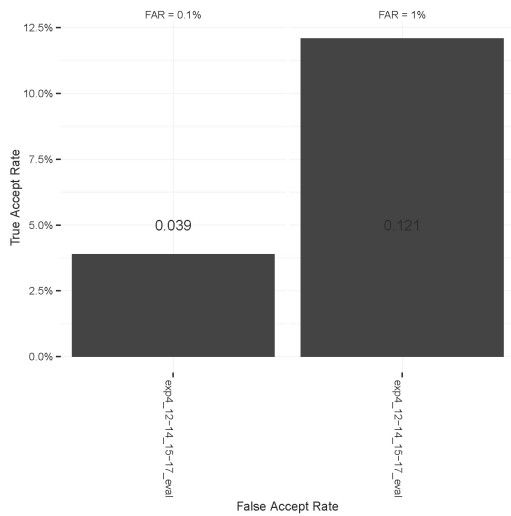
(b) S4F



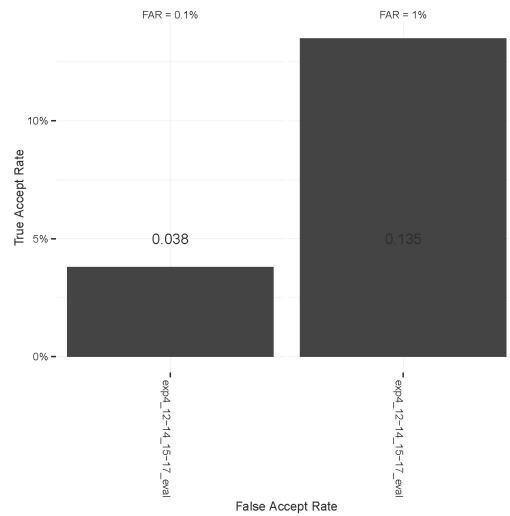
(c) CohortLDA



(d) LRPCA

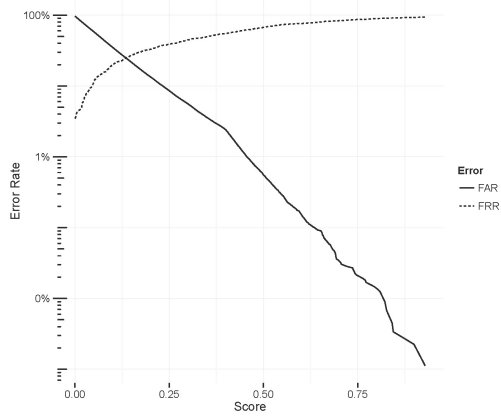


(e) LDA

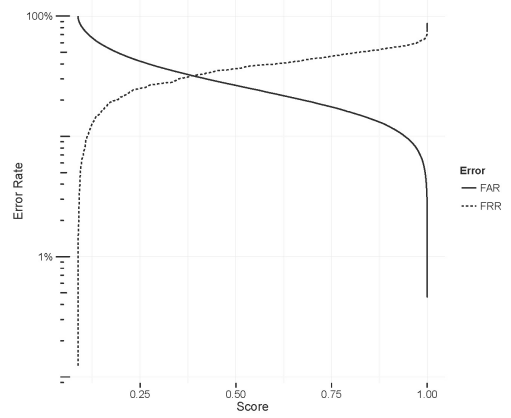


(f) PCA

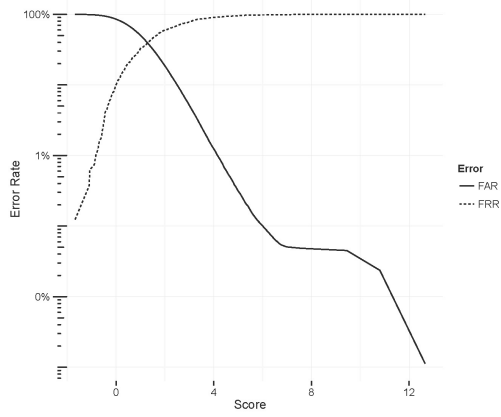
Figure 6.85: Experiment 4 12-14 to 15-17 - ROC Scores



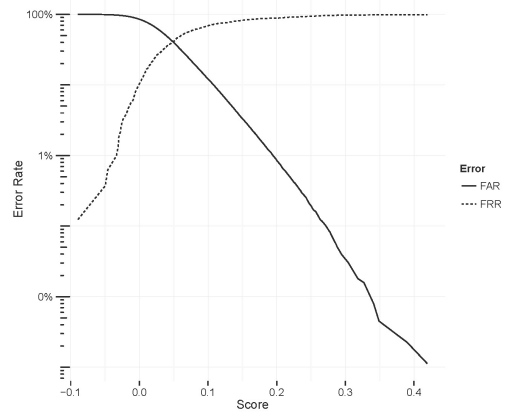
(a) Cognitec



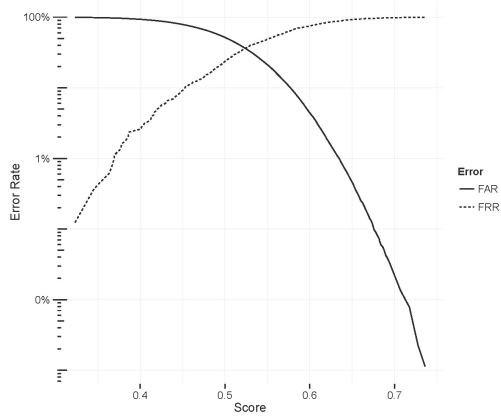
(b) S4F



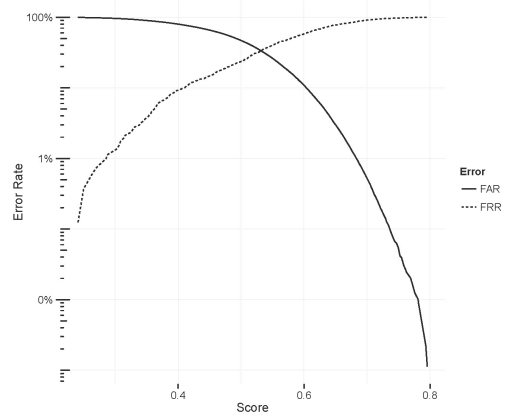
(c) CohortLDA



(d) LRPCA



(e) LDA

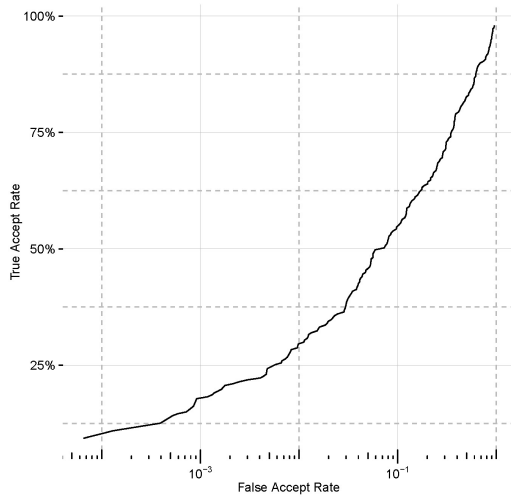


(f) PCA

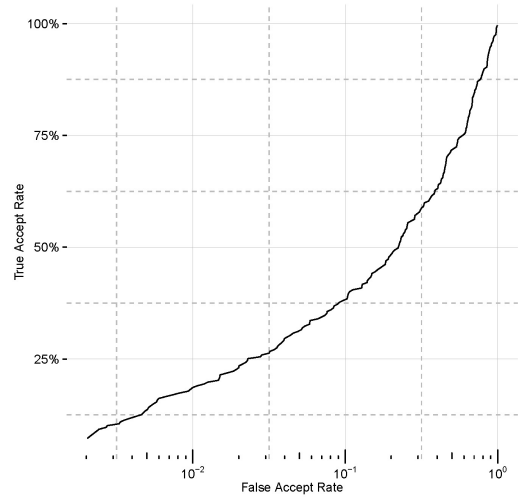
Figure 6.86: Experiment 4 12-14 to 15-17 - EER

APPENDICES M

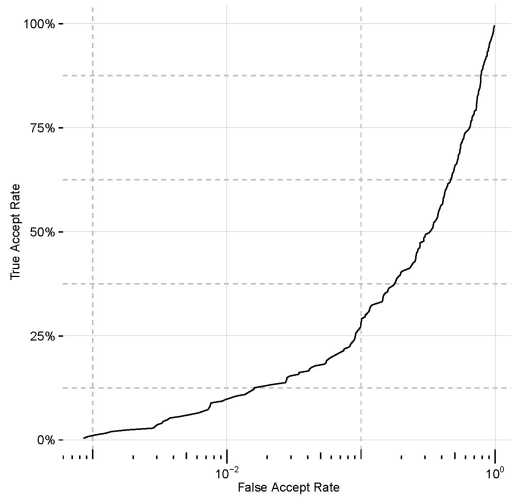
Experiment 4 - 12-14 to 18+ Results



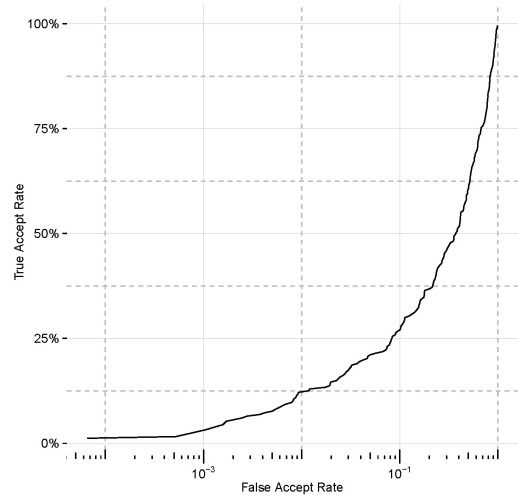
(a) Cognittec



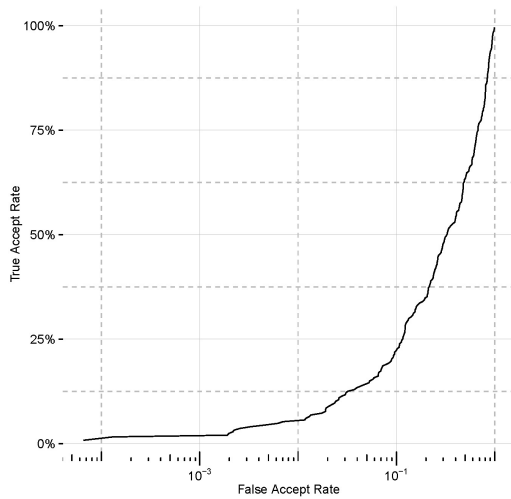
(b) S4F



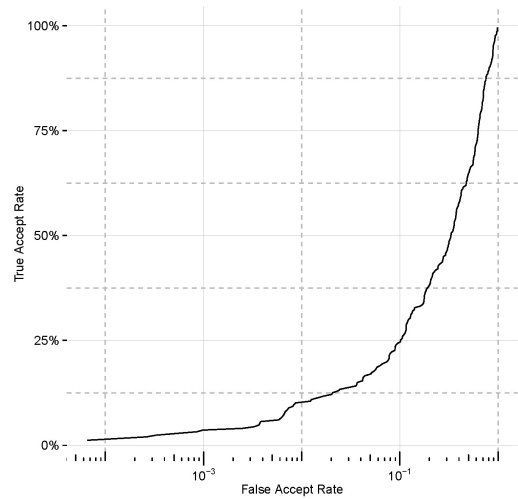
(c) CohortLDA



(d) LRPCA

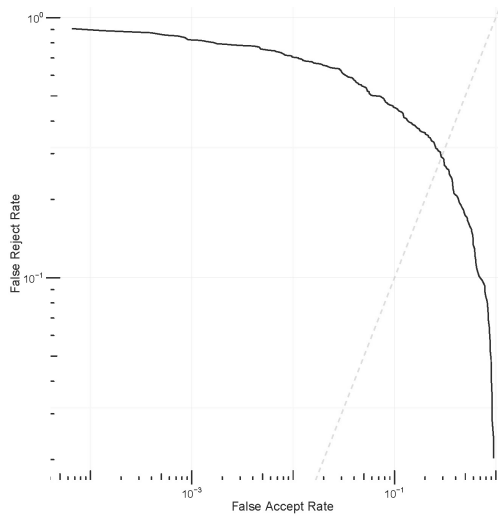


(e) LDA

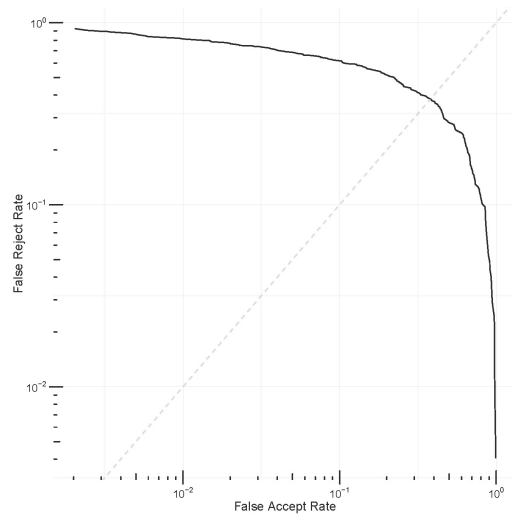


(f) PCA

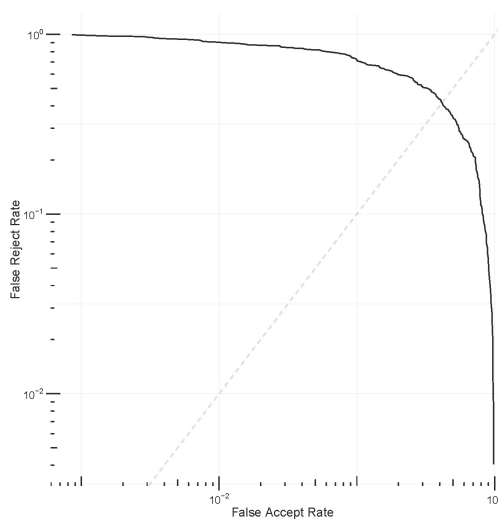
Figure 6.87: Experiment 4 12-14 to 18+ - ROC



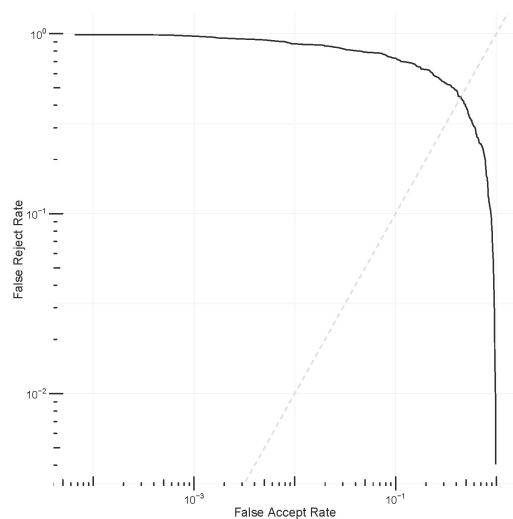
(a) Cognittec



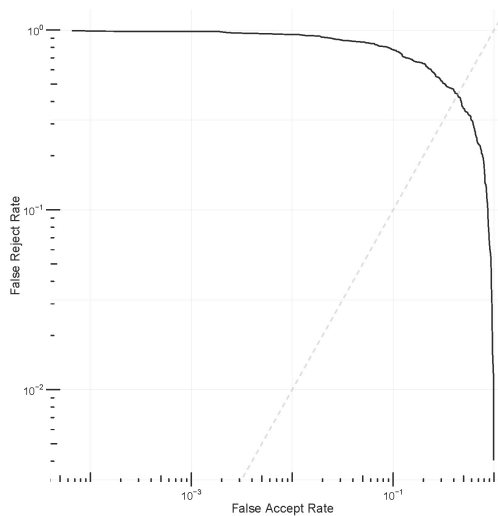
(b) S4F



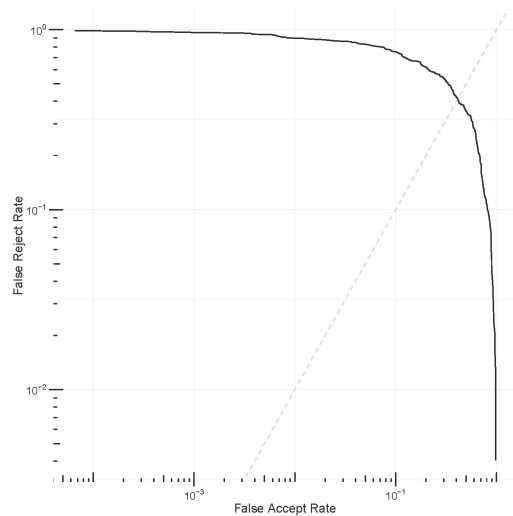
(c) CohortLDA



(d) LRPCA

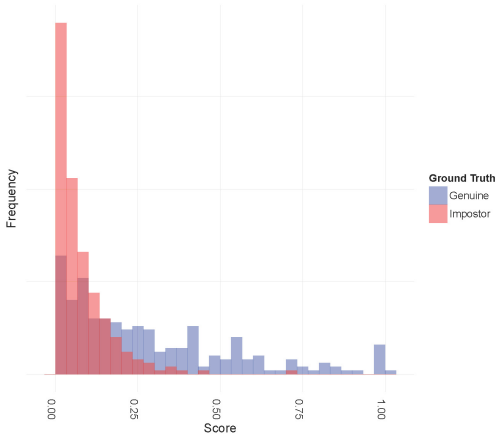


(e) LDA

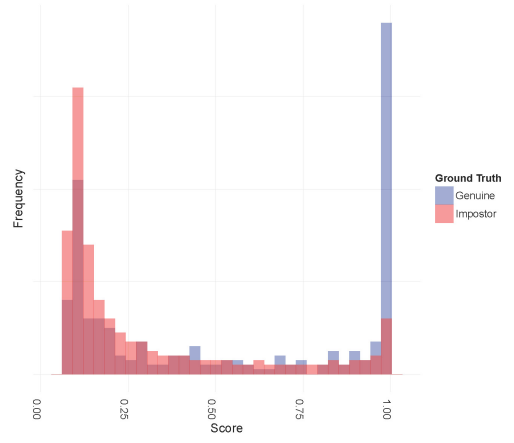


(f) PCA

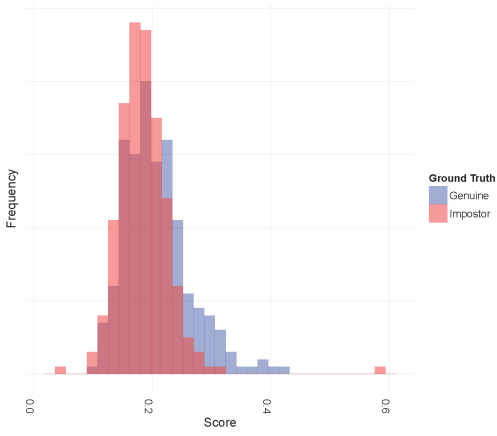
Figure 6.88: Experiment 4 12-14 to 18+ - DET



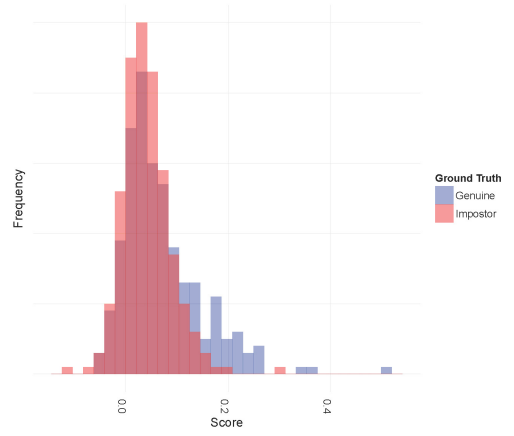
(a) Cognitec



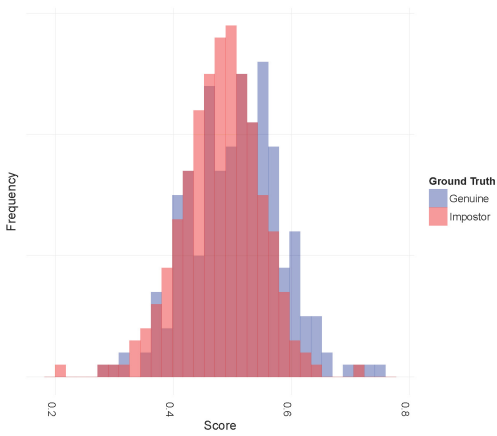
(b) S4F



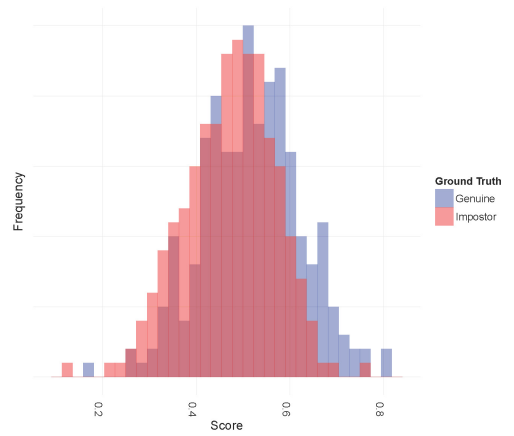
(c) CohortLDA



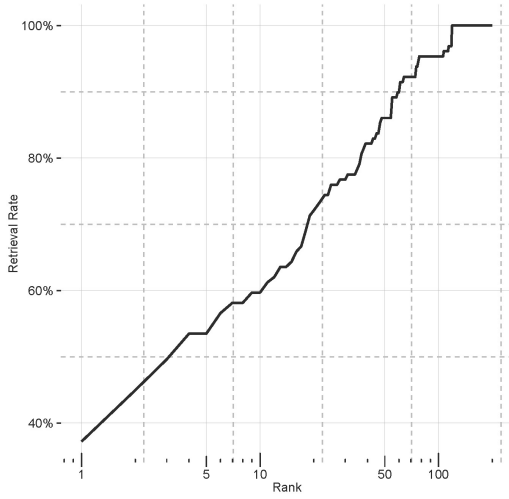
(d) LRPCA



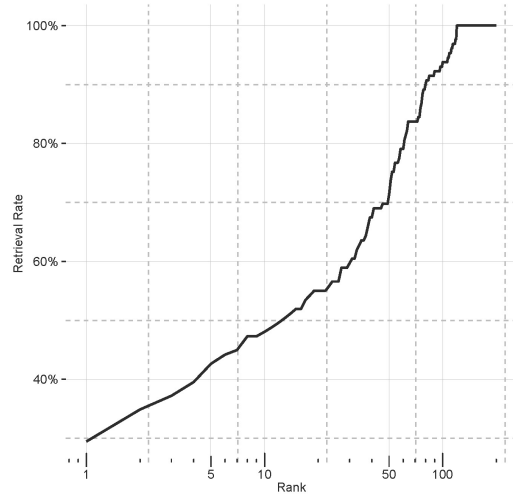
(e) LDA



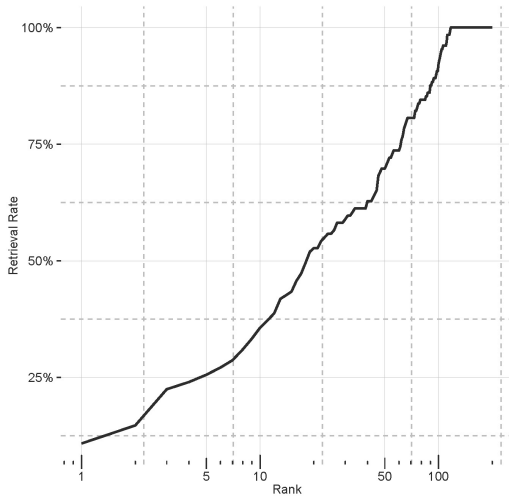
(f) PCA



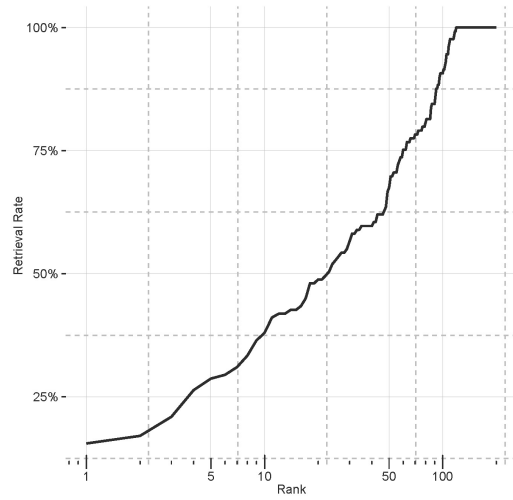
(a) Cognitec



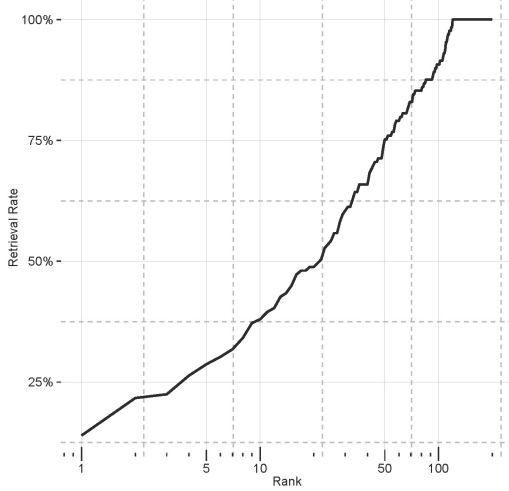
(b) S4F



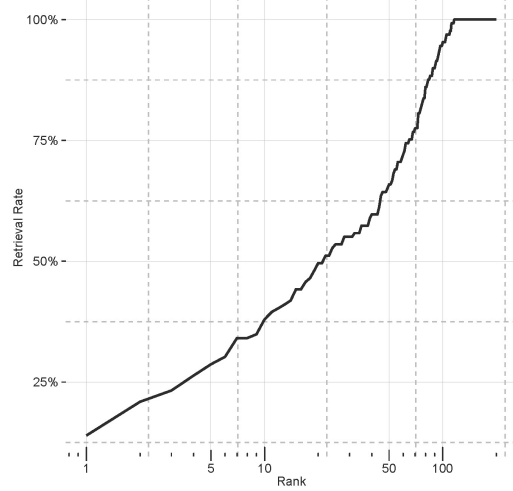
(c) CohortLDA



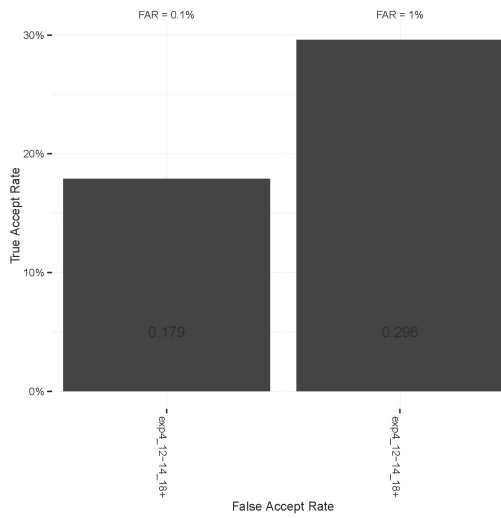
(d) LRPCA



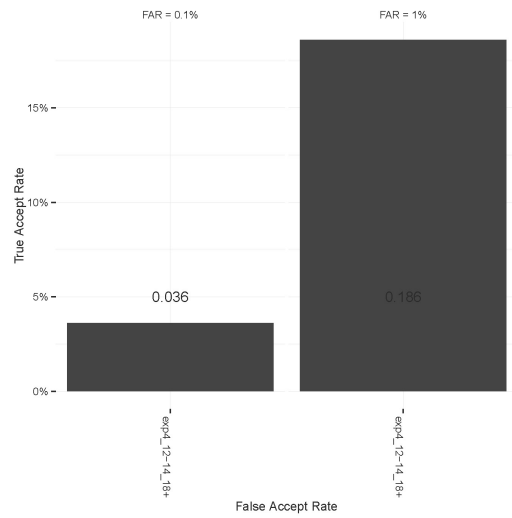
(e) LDA



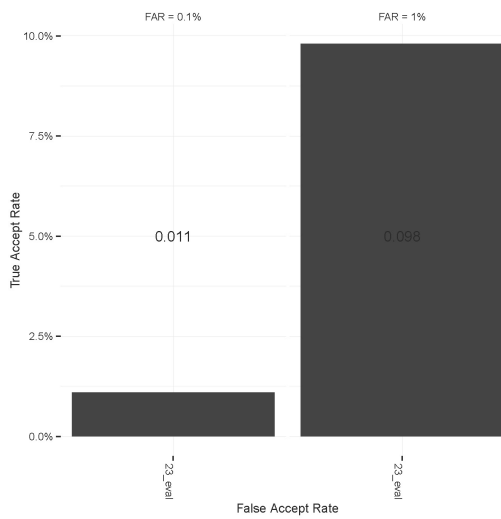
(f) PCA



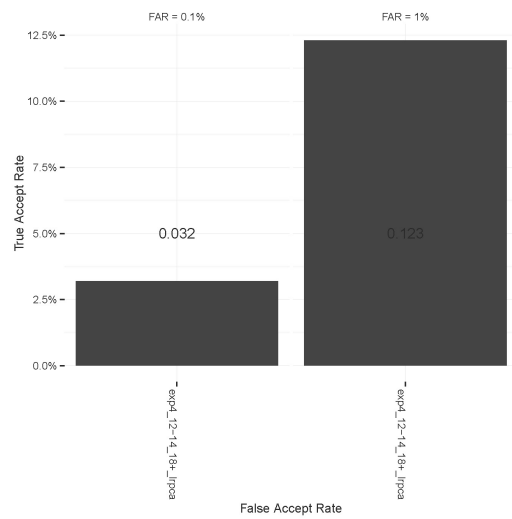
(a) Cognitec



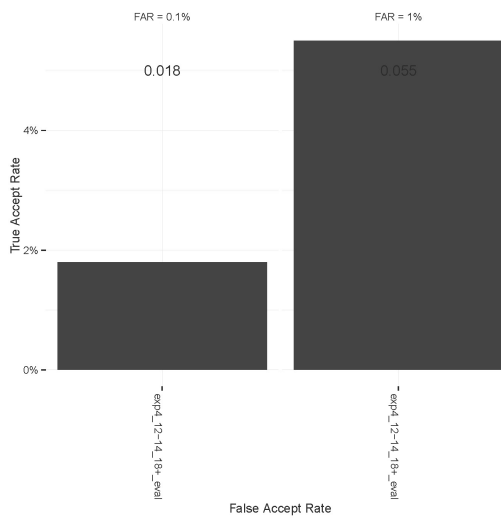
(b) S4F



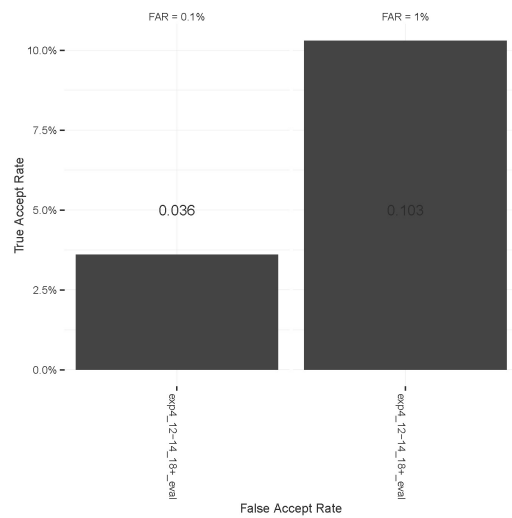
(c) CohortLDA



(d) LRPCA

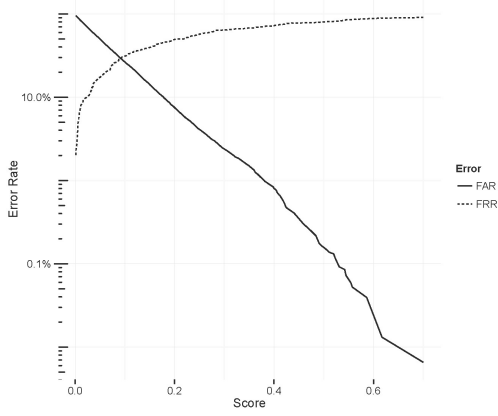


(e) LDA

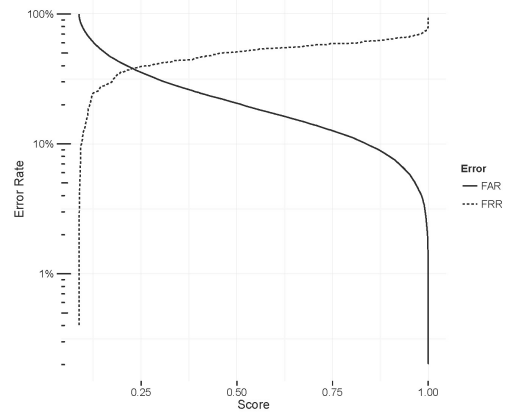


(f) PCA

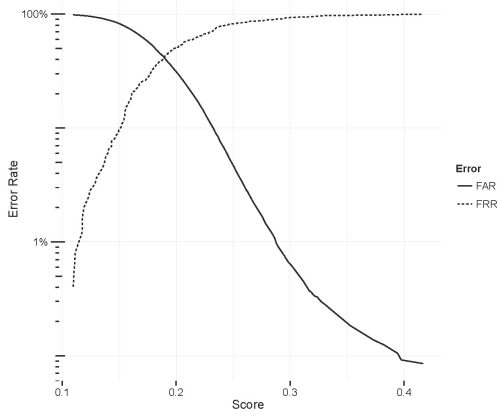
Figure 6.91: Experiment 4 12-14 to 18+ - ROC Scores



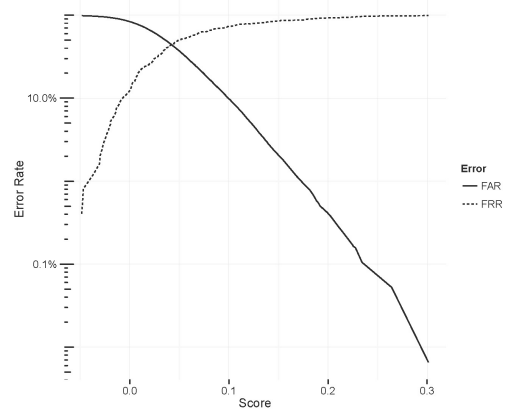
(a) Cognitec



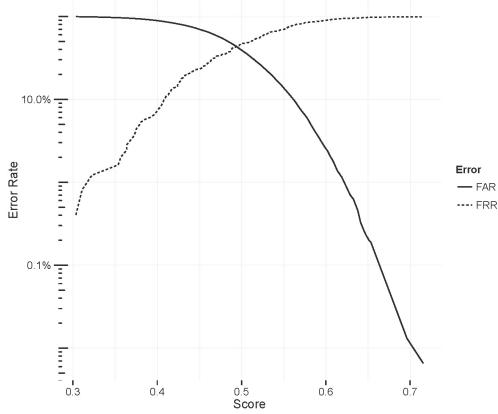
(b) S4F



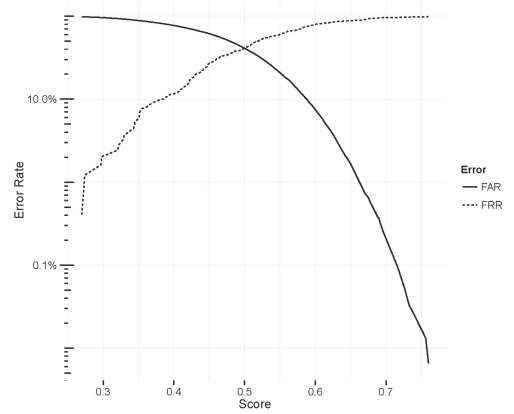
(c) CohortLDA



(d) LRPCA



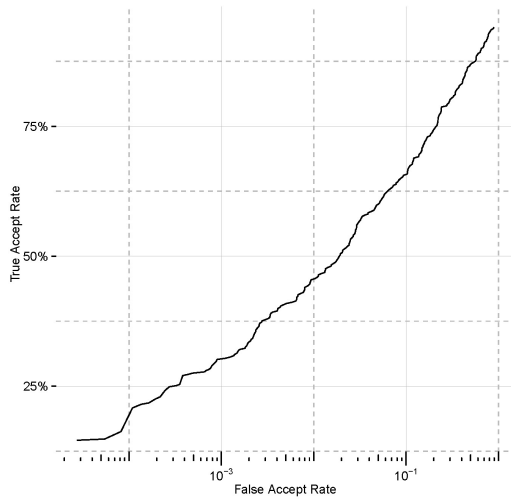
(e) LDA



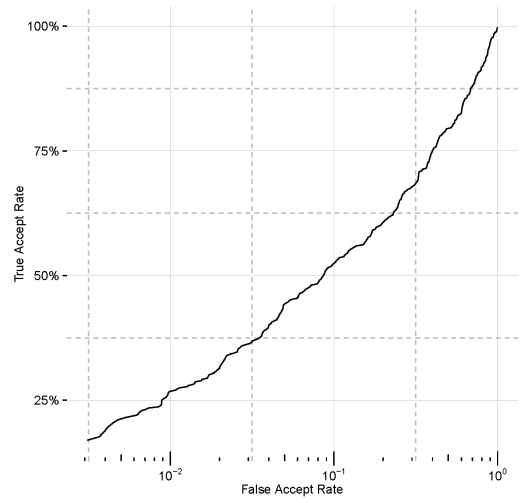
(f) PCA

APPENDICES N

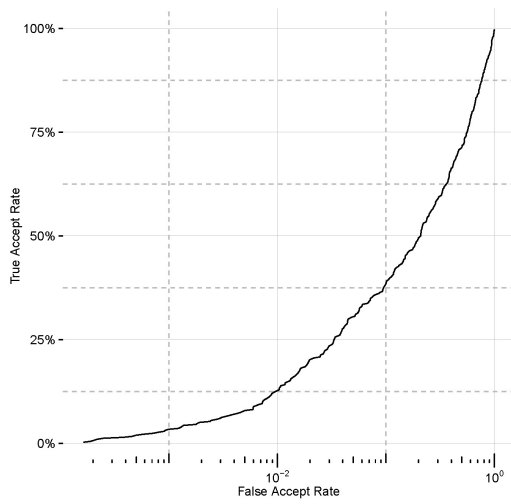
Experiment 4 - 15-17 to 18+ Results



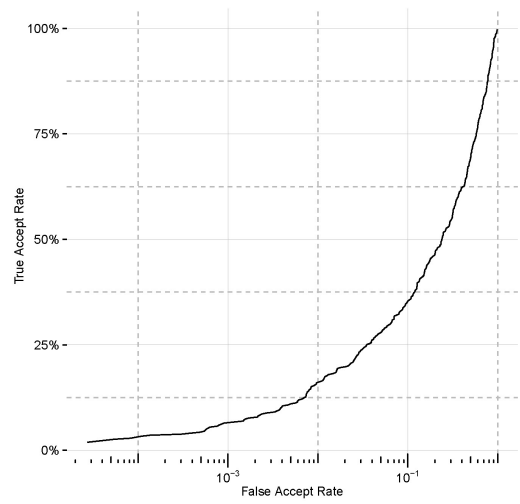
(a) Cognitec



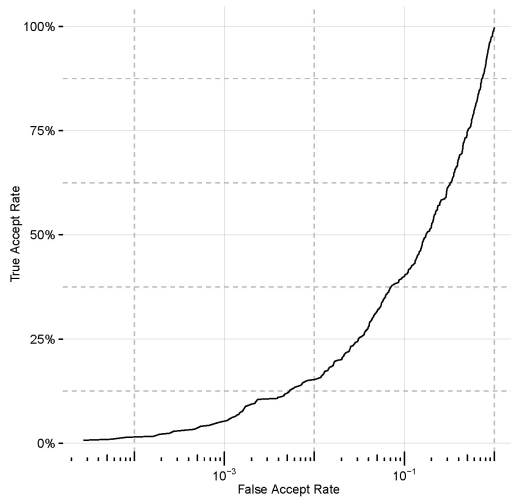
(b) S4F



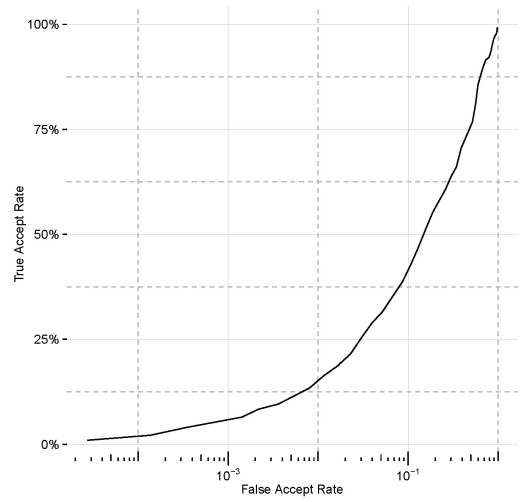
(c) CohortLDA



(d) LRPCA

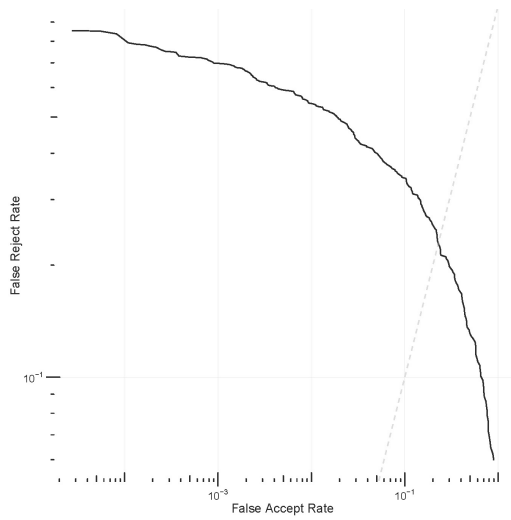


(e) LDA

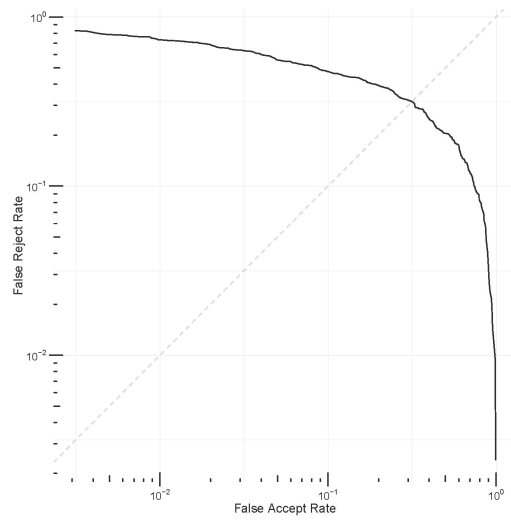


(f) PCA

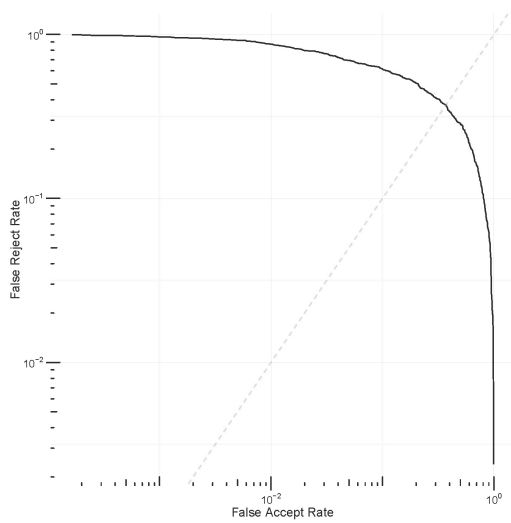
Figure 6.93: Experiment 4 15-17 to 18+ - ROC



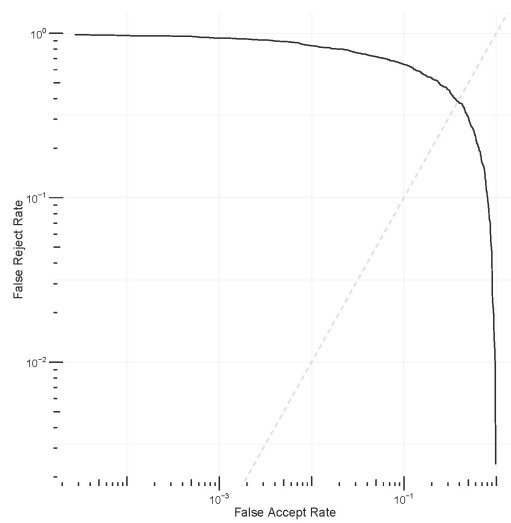
(a) Cognitec



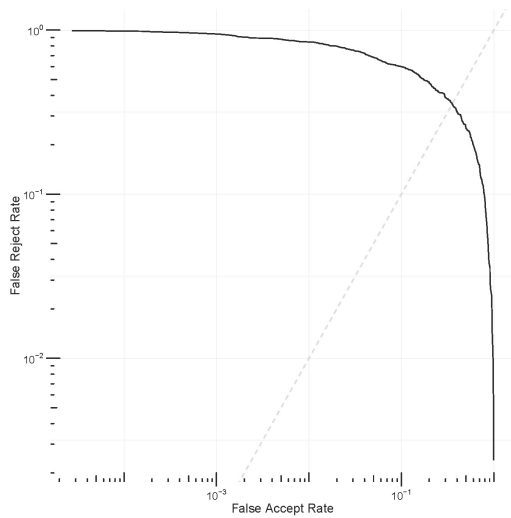
(b) S4F



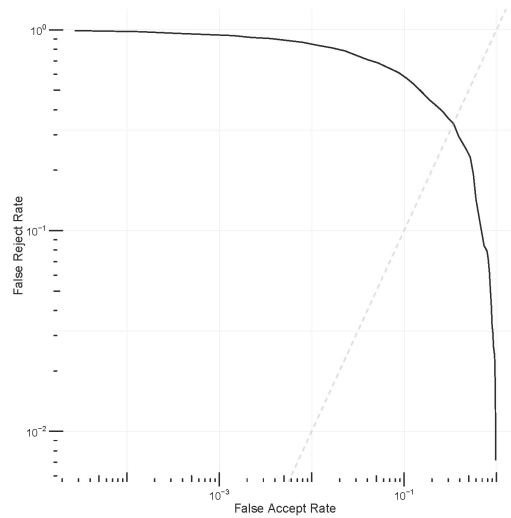
(c) CohortLDA



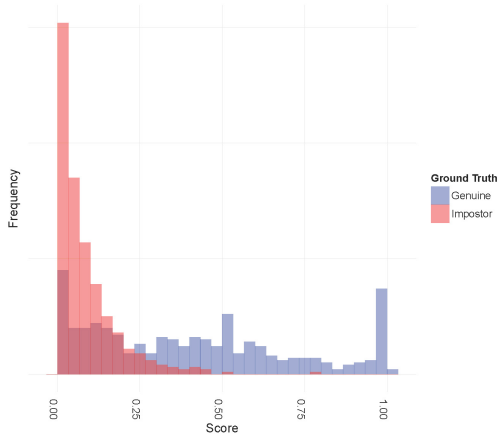
(d) LRPCA



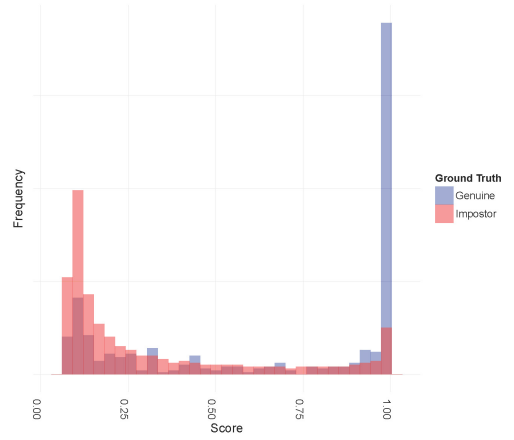
(e) LDA



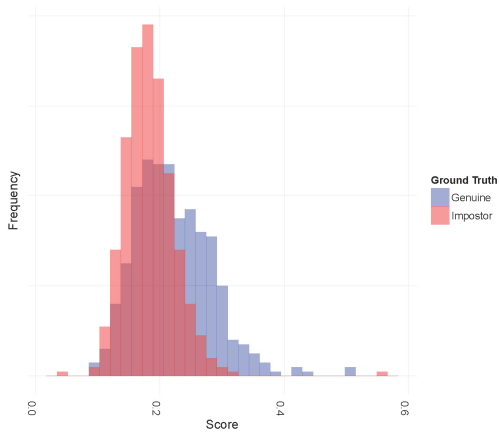
(f) PCA



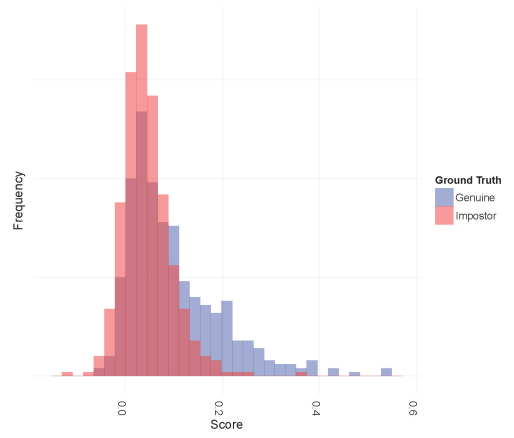
(a) Cognitec



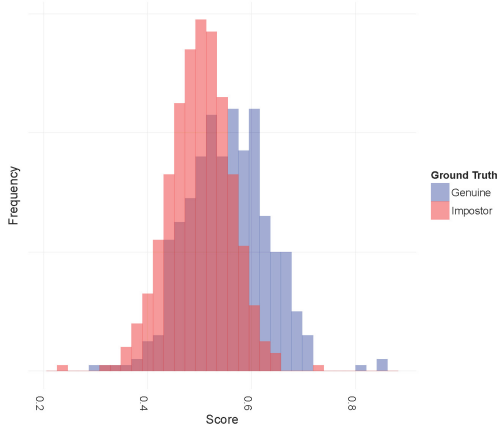
(b) S4F



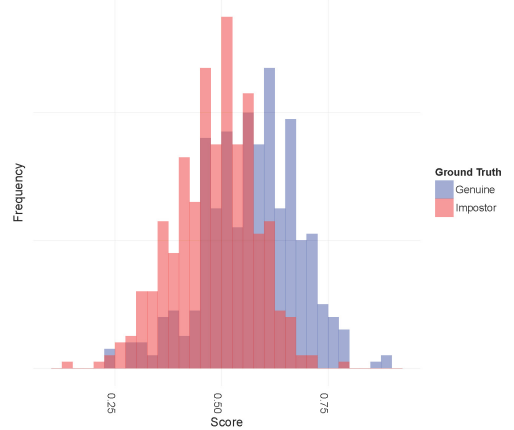
(c) CohortLDA



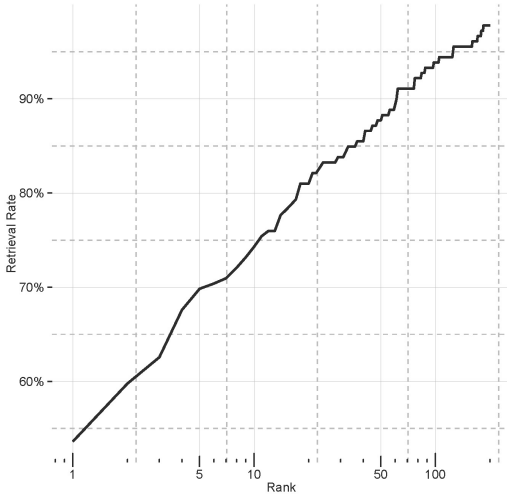
(d) LRPCA



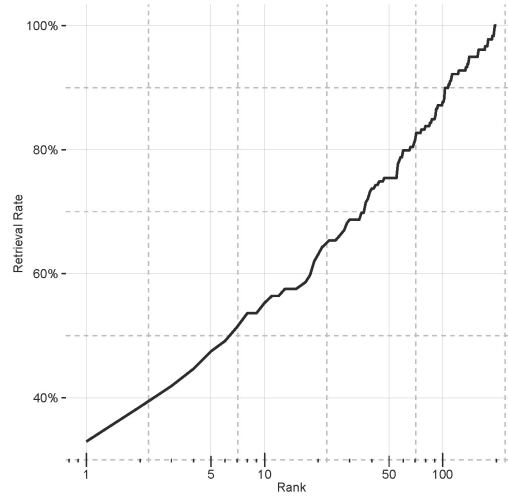
(e) LDA



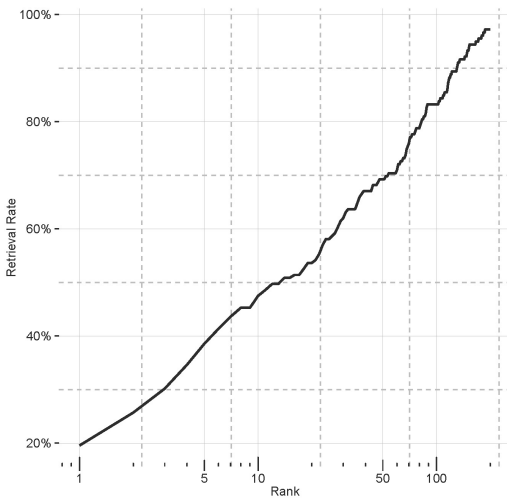
(f) PCA



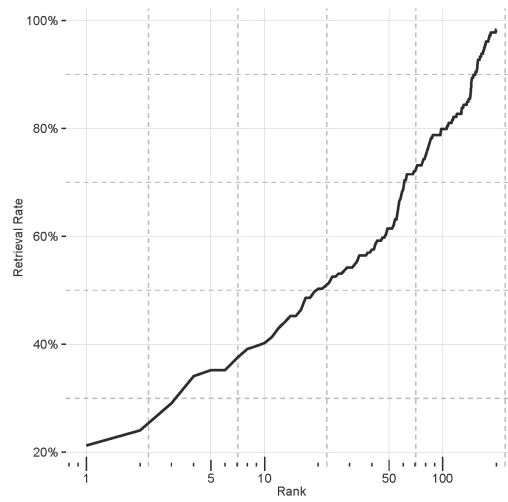
(a) Cognittec



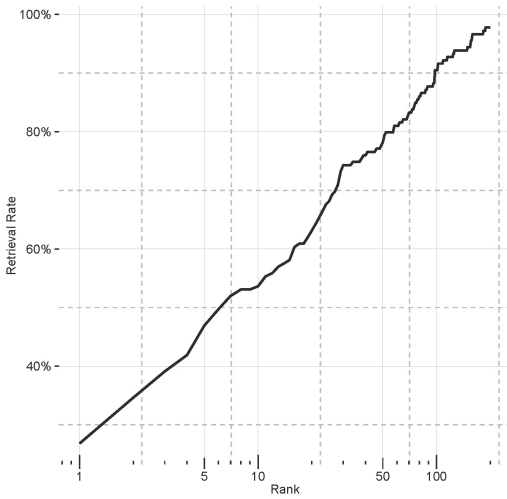
(b) S4F



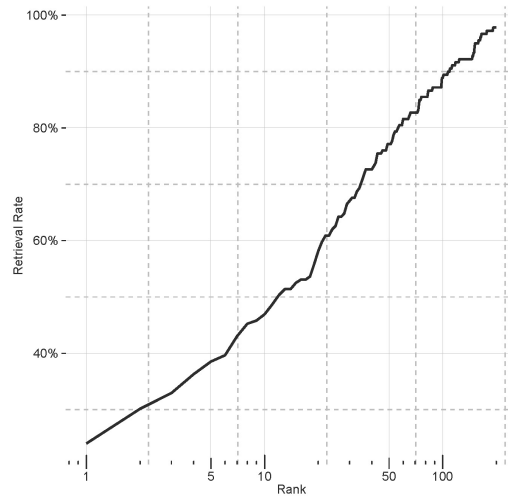
(c) CohortLDA



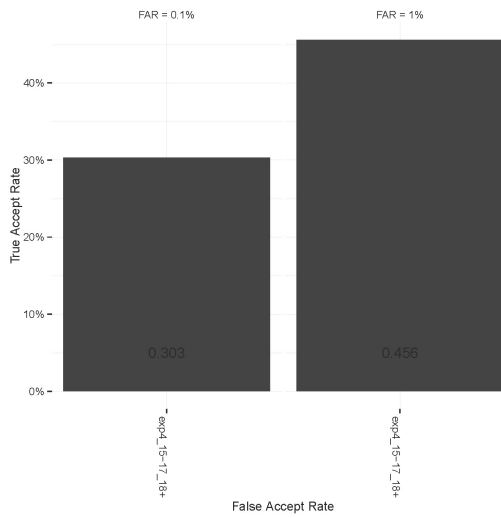
(d) LRPCA



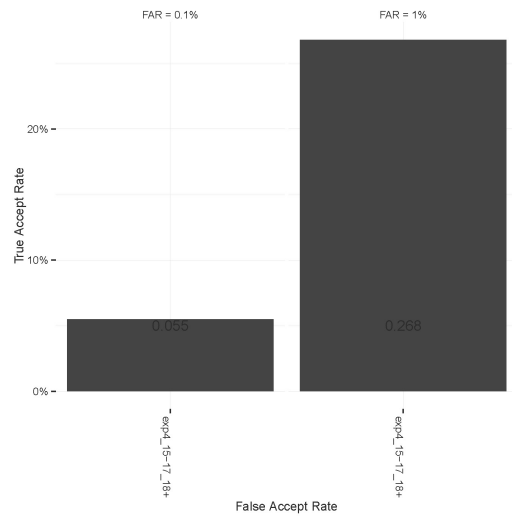
(e) LDA



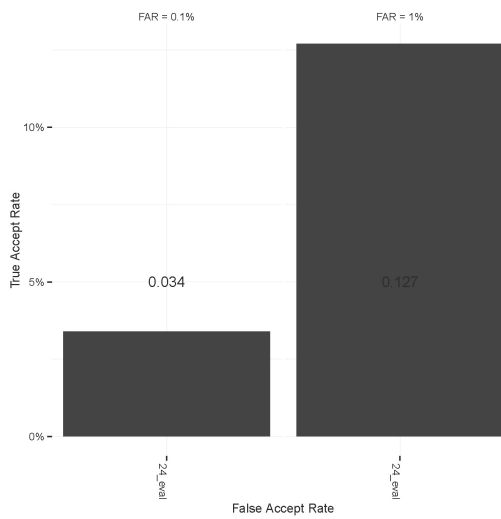
(f) PCA



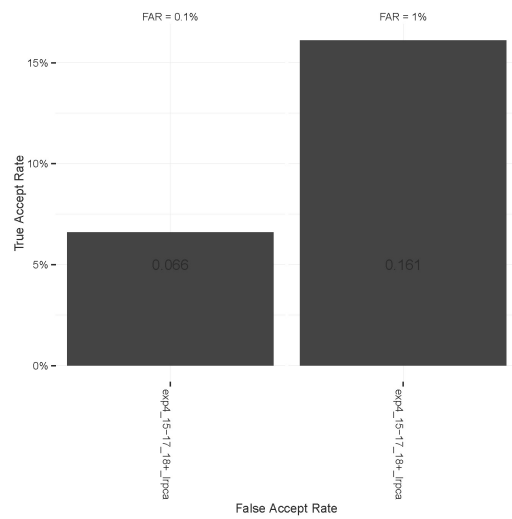
(a) Cognitec



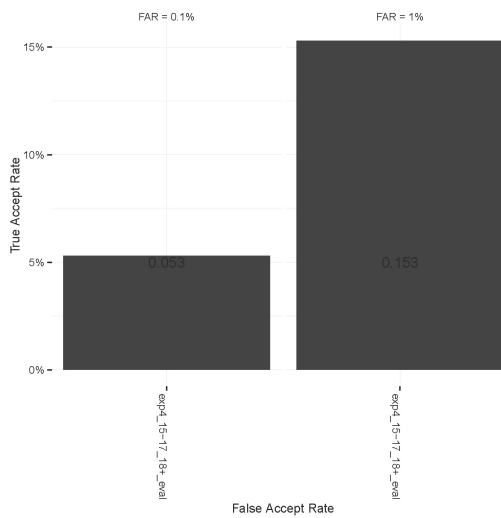
(b) S4F



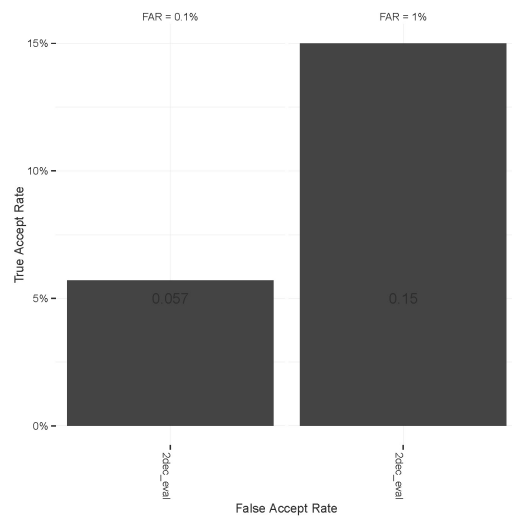
(c) CohortLDA



(d) LRPCA

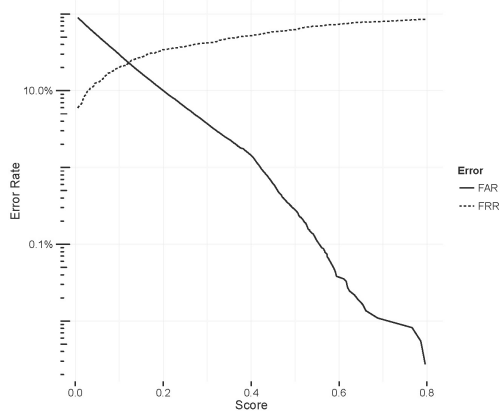


(e) LDA

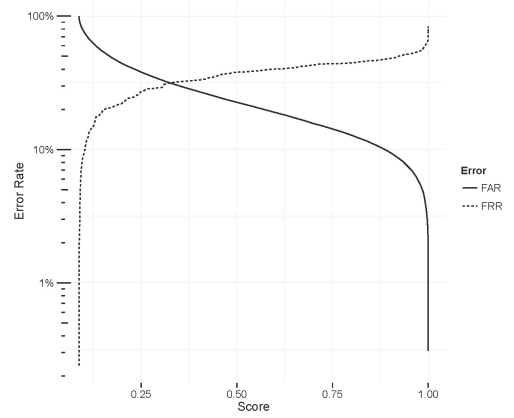


(f) PCA

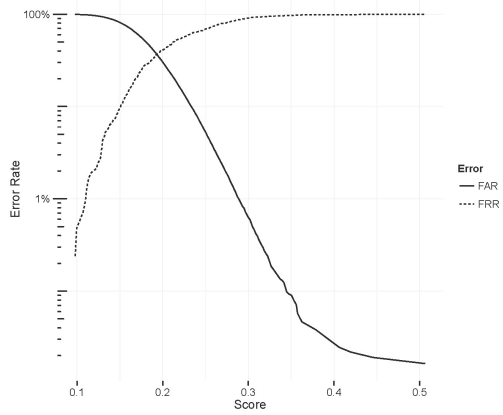
Figure 6.97: Experiment 4 15-17 to 18+ - ROC Scores



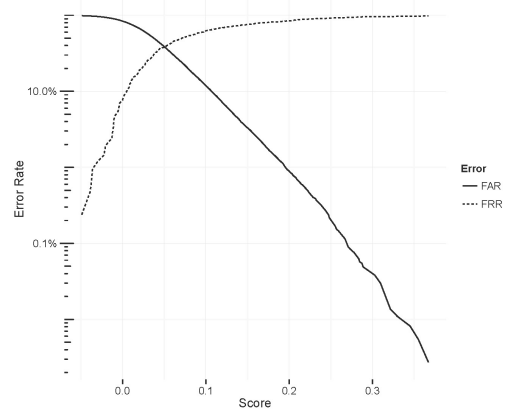
(a) Cognitec



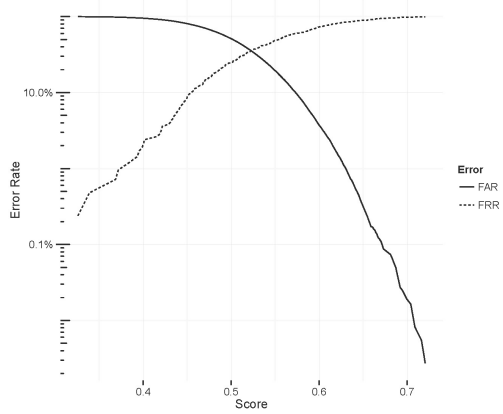
(b) S4F



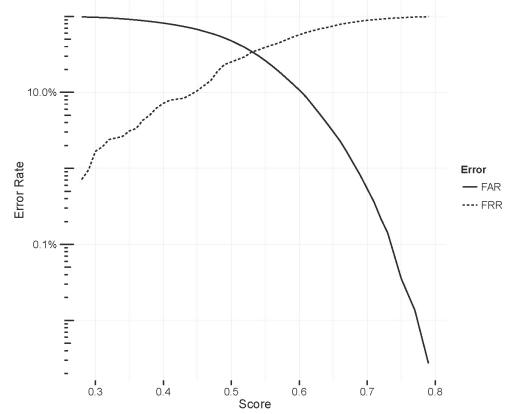
(c) CohortLDA



(d) LRPCA



(e) LDA



(f) PCA