AN INTELLIGENT MODEL FOR AGE AND GENDER PREDICTION FROM ARABIC TWITTER USER'S POSTS.

BY:

FAHAD AL MEKHLAFI

A THESIS

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF:

MASTER OF SCIENCE IN INFORMATION SECURITY

TO:

INFORMATION SECURITY DEPARTMENT

THE COLLEGE OF COMPUTER AND INFORMATION SECURITY

NAIF ARAB UNIVERSITY FOR SECURITY SCIENCES

2017G-1438H

SUPERVISED BY:

DR. HUSSEIN Y. MANSOUR

NAIF ARAB UNIVERSITY FOR SECURITY SCIENCES

Table of Contents

List of Figures

List of Tables

List of Acronyms

Acronyms	Details
ANN	Artificial Neural Networks.
URL	Uniform Resource Locator.
KDD	Knowledge Discovery and Data Mining
AI	Artificial Intelligent.
ML	Machine Learning.
DL	Deep Learning.
DM	Data Mining.
KNN	K Nearest Neighbors.
SVM	Support Vector Machines
SVN	Support Vector Network.
LIBSVM	Library for Support Vector Machine.
OVA	One-Vs-All.
OVO	One-Vs-One.
MDs	Multilabel Datasets.
MC	Multilabel Classification.
MLP	Multilabel Prediction.
DT	Data Transformations.
BR	Binary Relevance.
LP	Label PowerSet.
RPC	Ranking by Pairwise Comparison.
CLR	Calibrated Label Ranking.
MEKA	Multi-label Extension to WEKA.

Acronyms	Details
WEKA	Waikato Environment for Knowledge Analysis.
LC	Label Calibration.
RT	Ranking and Threshold.
TC	Text Classification.
NB	Naïve Bays
ML	Multi Label
MFCC	Mel-Frequency Cepstral Coefficients.
SFM	Spectral Flatness Measure.
RBFKernel	Radial Basis Function Kernel.
LIBINAR	Library for Large Linear Classification
TF-IDF	Term Frequency-Inverse Document Frequency
HTML	Hyper Text Markup Language
Ws	Slang Word.
LIWC	Words Linguistic Inquiry and Word Count dictionary.
BoW	Bag of words.
TF	Term Frequency.
TO	Term Occurrences.
BTO	Binary Term Occurrences.
VSM	Vector Space Model.
ARI	Automated Readability Index.
CLI	The Coleman-Liau Index.
RIX	The Rix Readability Index.
NLP	Natural Languages Processing.

Acronyms	Details
LSI	Latent Semantic Indexing.
LPI	Locality Preserving Indexing.