SENTIMENT ANALYSIS AND EMOTION DETECTION ON CRYPTOCURRENCY RELATED TWEETS USING ENSEMBLE LSTM–GRU MODEL

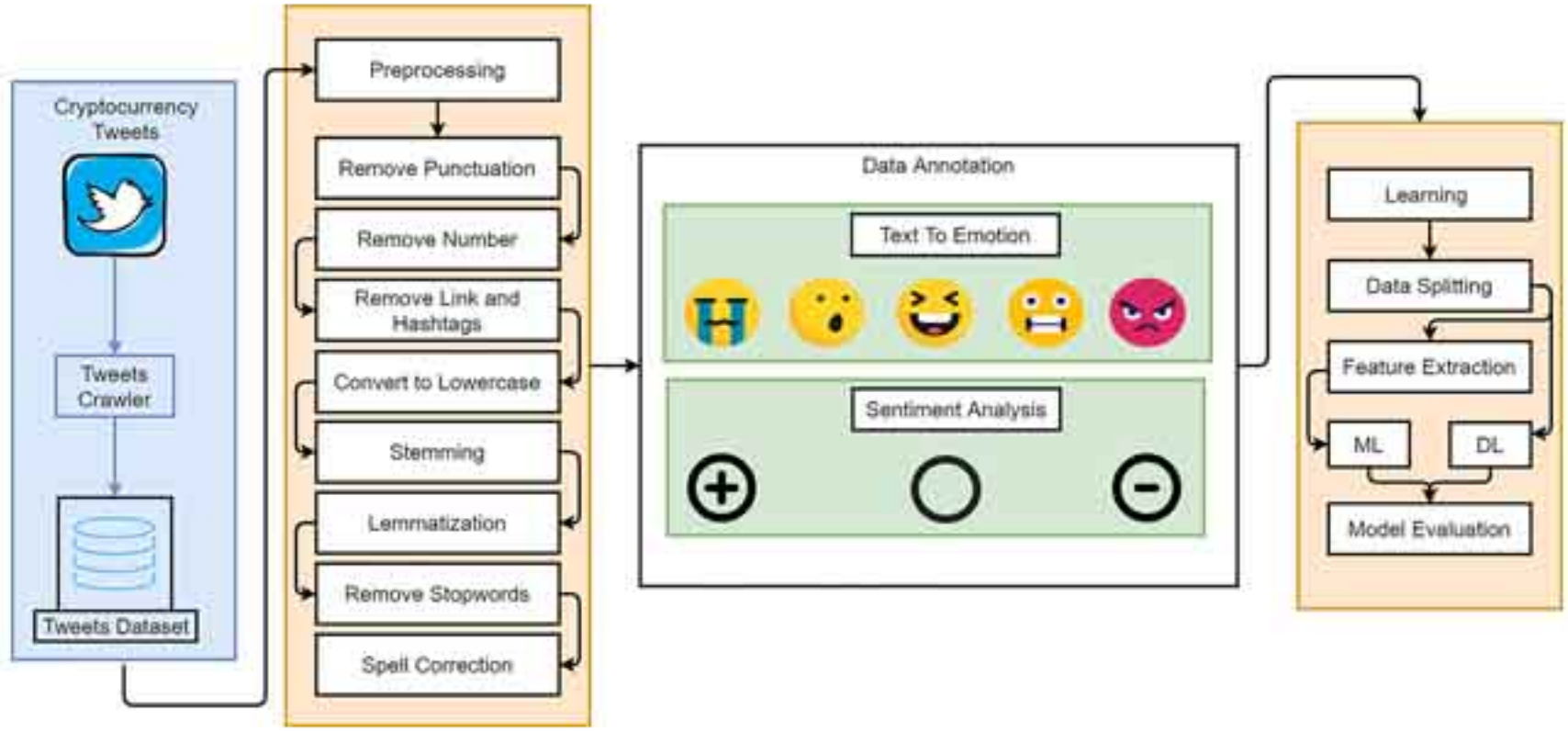
Presented By : Moumita Khandaker
ID : 21266013
CSE707 Task 2

CONTENTS

INTRODUCTION

Cryptocurrency sentiment analysis is potentially important since it is frequently used to forecast market prices, which requires highly accurate sentiment classification. This study uses tweets about cryptocurrencies, which are commonly used to predict their market prices, to perform sentiment analysis and emotion detection. Long short term memory and gated recurrent unit, two recurrent neural network applications, are combined in the deep learning ensemble model LSTM-GRU to increase the analysis's efficacy.

PROPOSED METHODOLOGY



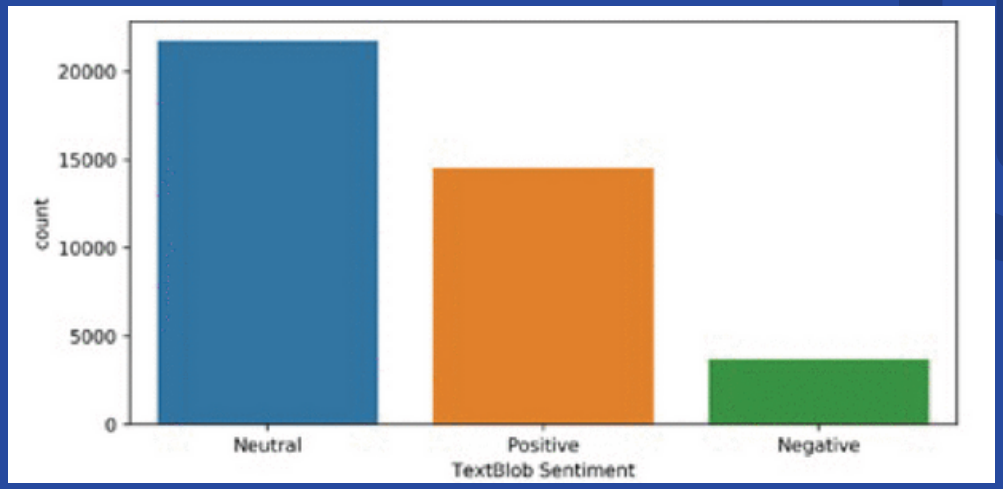
Architecture

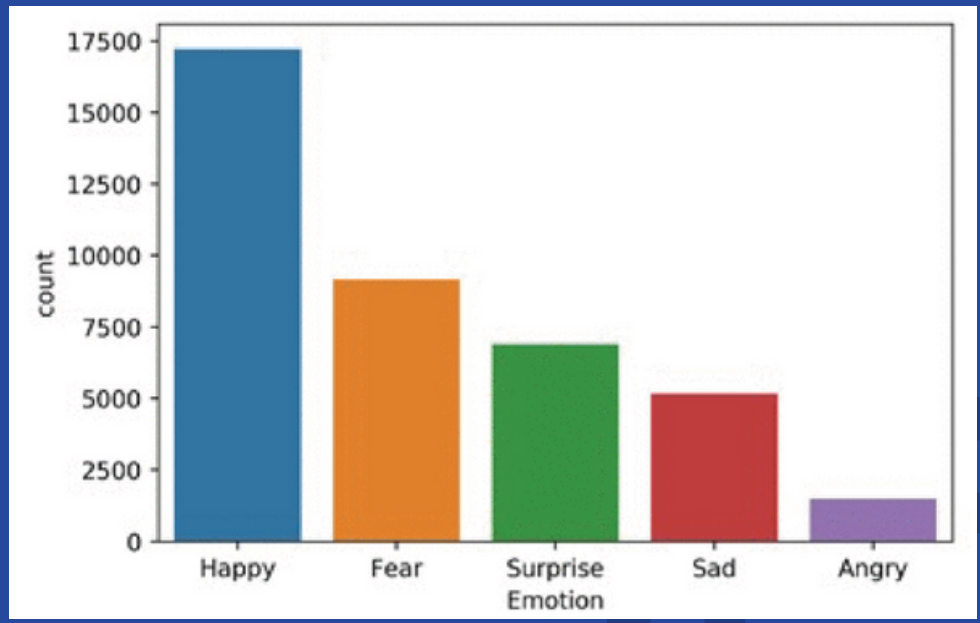
PROPOSED METHODOLOGY

Text	TextBlob		Text2Emotion	
	Polarity score	Sentiment	Highest score	Emotion
nice trad bot just sold neo usdt profit on bi-nance binance fee	0.6	Positive	0.5	Happy
striatu world first unreinforc print concrete archi bridge blockchain cryptocurrency	0.2	Positive	0.5	Surprise

Sentiment Scoring Using TextBlob and Text2Emotion

Proposed Methodology





NUMBER OF SAMPLES, EACH CLASS IN THE DATASET, EMOTIONS FOUND IN THE DATASET

NUMBER OF SAMPLES FOR TRAINING AND TEST DATASETS

34,000 Training Set

6,000 Testing Set

40,000 Total Data

PREPROCESSING

1. Tokenization
2. Punctuation Removal
3. Number Removal
4. Stemming
5. Lemmatization
6. Stop Words Removal
7. Spell Correction

FEATURE EXTRACTION

1. Bag of Words
2. Term Frequency-Inverse Document Frequency
3. Word2Vec

EMOTION DETECTION AND SENTIMENT ANALYSIS METHODS

1 TextBlob

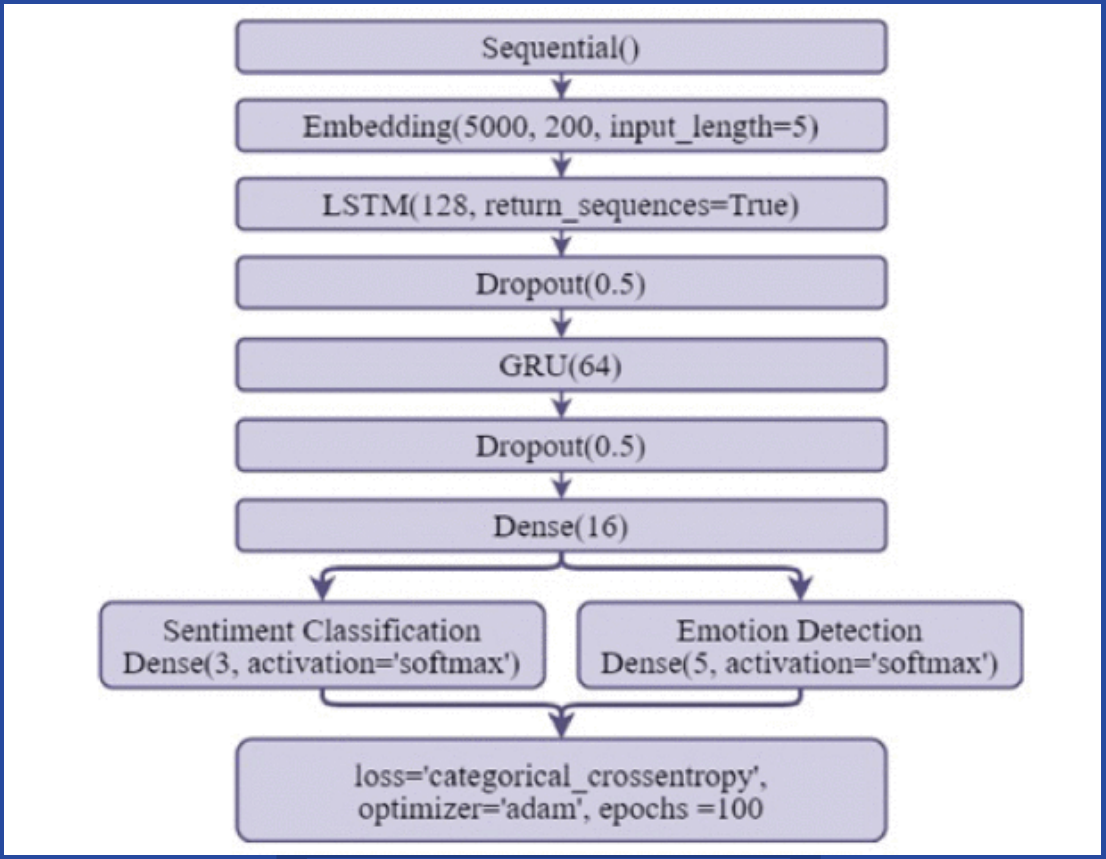
2 Text2Emotion

MACHINE LEARNING MODELS

Model	Hyperparameters	Tuning range
Random Forest	n_estimators= 300; max_depth=300	n_estimators= 50 to 500; max_depth=50 to 500;
Decision Tree	max_depth=300	n_estimators= 50 to 500;
K Nearest Neighbour	n_neighbors= 3; weights=uniform	n_neighbors= 1 to 5; weights= {uniform, distance}
Support Vector Machine	kernel= linear; C=3.0	kernel= {linear, poly}; C=1.0 to 5.0
Gaussian Naive Bayes	var_smoothing= 1e-9	var_smoothing= 1e-9
Logistic Regression	solver= liblinear; C=3.0	solver= {liblinear, saga, sag}; C=1.0 to 5.0

Hyper-Parameters Setting for Machine Learning Models

PROPOSED ENSEMBLE MODEL



Structure of the proposed deep learning based ensemble model

Results and Discussions

EMOTION DETECTION RESULTS USING MACHINE LEARNING MODELS

Model	Accuracy	Precision	Recall	F1 score	G mean
SVM	0.90	0.89	0.89	0.89	0.89
LR	0.90	0.90	0.88	0.89	0.89
GNB	0.38	0.41	0.44	0.35	0.42
ETC	0.85	0.88	0.77	0.81	0.82
DT	0.84	0.83	0.79	0.80	0.82
KNN	0.69	0.81	0.56	0.63	0.67

Results Using Machine Learning Models with BoW Features

Model	Accuracy	Precision	Recall	F1 score	G mean
SVM	0.90	0.88	0.86	0.87	0.87
LR	0.90	0.90	0.84	0.87	0.87
GNB	0.37	0.40	0.44	0.35	0.42
ETC	0.84	0.89	0.75	0.80	0.82
DT	0.83	0.81	0.77	0.79	0.79
KNN	0.69	0.81	0.56	0.63	0.67

Results Using Machine Learning Models with TF-IDF Features

Results and Discussions

EMOTION DETECTION RESULTS USING MACHINE LEARNING MODELS

Model	Accuracy	Precision	Recall	F1 score	G mean
SVM	0.75	0.74	0.67	0.70	0.70
LR	0.76	0.74	0.68	0.70	0.71
GNB	0.42	0.47	0.43	0.37	0.45
ETC	0.74	0.87	0.60	0.67	0.72
DT	0.63	0.55	0.55	0.55	0.55
KNN	0.72	0.74	0.61	0.66	0.67

Results Using Machine Learning Models with Word2Vec Features

Model	BoW		TF-IDF		Word2Vec	
	CP	WP	CP	WP	CP	WP
SVM	5,417	583	5,389	611	4,514	1,486
LR	5,406	594	5,375	625	4,539	1,461
GNB	2,251	3,749	2,243	3,757	2,526	3,474
ETC	5,102	898	5,054	946	4,435	1,565
DT	5,048	952	4,964	1,036	3,784	2,216
KNN	4,117	1,883	4,117	1,883	4,315	1,685

Results Comparison Using Machine Learning Models

DRAWBACK

When it comes to right and incorrect predictions for sentiment analysis and emotion detection, LSTM-GRU performs better than any other model. The utilisation of random undersampling for dataset balancing implies that a decrease in training data leads to a decrease in LSTM-GRU performance.

FUTURE WORK

Sentiment analysis and emotion detection are applied to cryptocurrency-related tweets in this study. Cryptocurrency sentiment analysis is potentially important since it is frequently used to forecast market prices, which requires highly accurate sentiment classification.

THANK YOU!