DEEP LEARNING FOR TRUTHFULNESS ASSESSMENT: DETECTING FAKE NEWS IN SOCIAL MEDIA THROUGH DEEP LEARNING

THESIS STUDENTS

MAHABUBUL ALAM
20101105
AUNANNA BINTE WAHID
20201167
MD RAKIBUL ISLAM
24341314
NABIL FAIEAZ DIPTA
20201180
TANMIN ALAM ROKTI
21101051

Supervisor

Dr. Muhammad Iqbal Hossain
Associate professor
Dept. of CSE
Brac University

THESIS ID: T23310114

INTRODUCTION

01

SPREAD OF MISINFORMATION IN SOCIAL MEDIA

02

IMPACT OF FAKE NEWS IN SOCIETY,POLITICS & PERSONAL LIFE

03

USE OF DEEP LEARNING TO DETECT FAKE NEWS

RESEARCH OBJECTIVES

01

EXPLORING ROBUST FAKE NEWS DETECTION MODELS

02

INVESTIGATING THE IMPACT OF HYBRID DEEP LEARNING MODELS

03

ANALYSING THE POTENTIAL OF DL MODELS FOR AUTOMATED FAKE NEWS DE TECTION

04

DIFFERENTIATING THE PERFORMANCE OF VARIOUS DEEP LEARNING MODELS

RELATED WORKS

Dataset Type	Model Name	Performance
BuzzFeed and PolitiFact,GeoText Dataset,UTGeo2011 Dataset	GCN, MENET	BuzzFeed accuracy-GCN :87.8%, Poli-tiFact accuracy-GCN: 89.5%
Custom News Dataset(Text based)	DNN	Accuracy: 91%
DataHeart dataset	LSTM, 14-Layer BLSTM	Acc: 91.08%, Rec: 91.10%, F1 : 91.57%
Snopes, PolitiFact, NewsTrust, SemEval-2017	DeClarE	DeClarE(Full) on Politifact Acc:67.32%, False claim Acc:69.62%, Macro F1 score:68%, AUC:75%

DATASET COLLECTION

Dataset Used

-CIC Truth Seeker Dataset 2023

ISOT Fake News Dataset

Data Collected Process

-Extracted tweets and articles from online sources

-Labeled data into 3-label and 5-label classification

Data Sources

TRUTH SEEKER DATASET:

-Collected tweets from PolitiFact Dataset

-Contains 186,000 tweets

ISOT Dataset:

-Contains 45,000 news articles

Dataset Features

-Textual-based

-Includes author metadata, tweet content, source credibility

DATA PREPROCESSING

Data Loading & Initial Cleaning

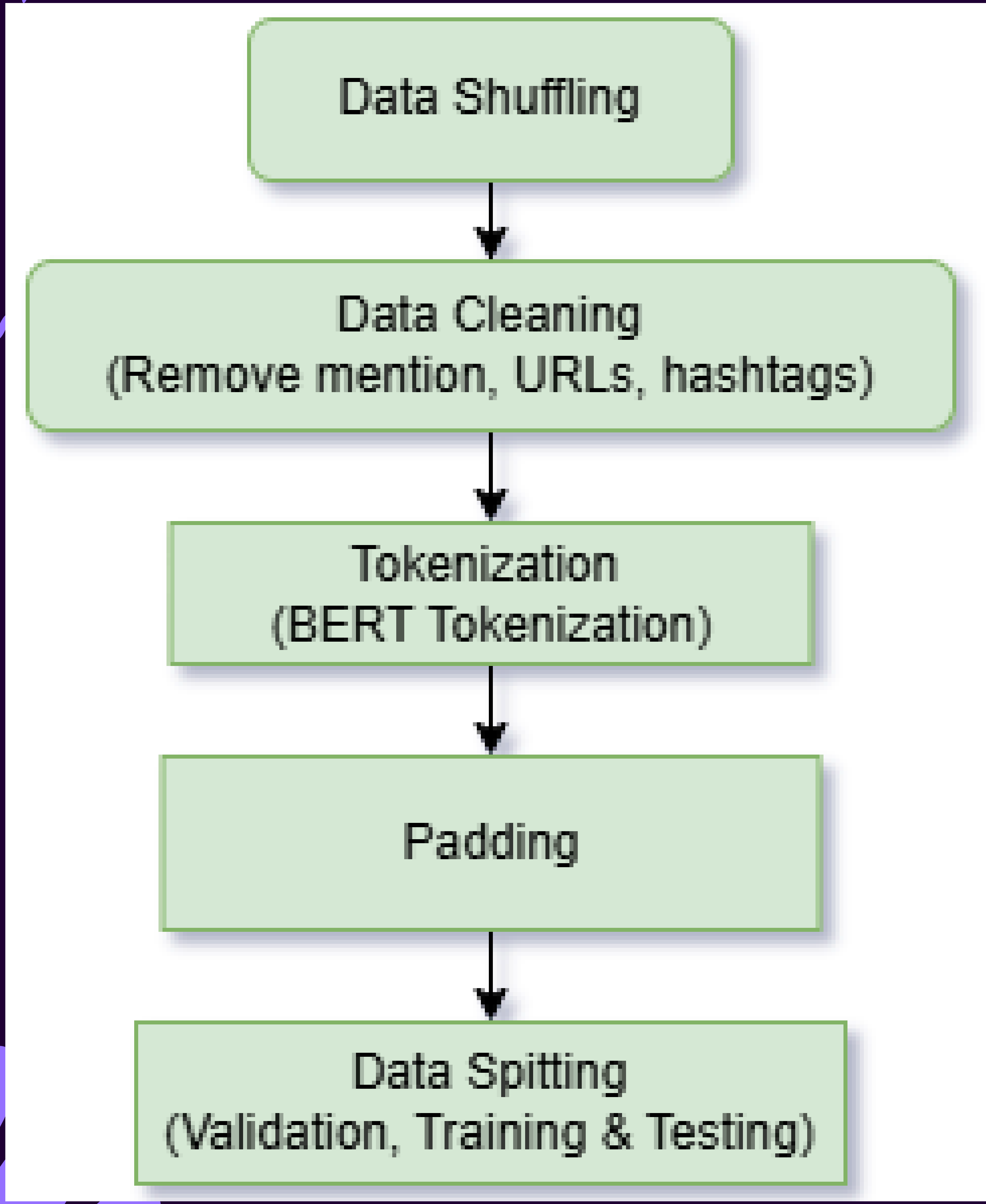
Data Shuffling

Text Cleaning & Normalization

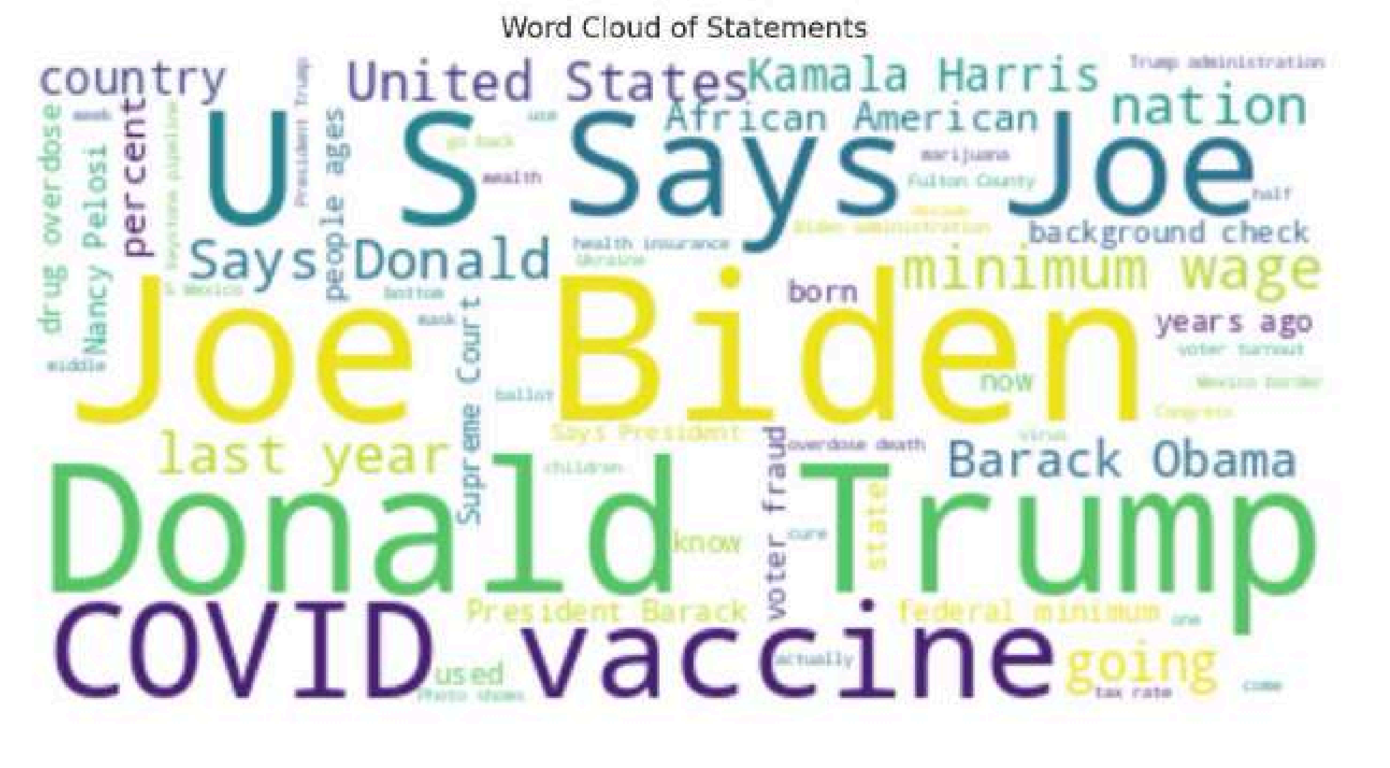
Tokenization & vectorization

Sequence Padding

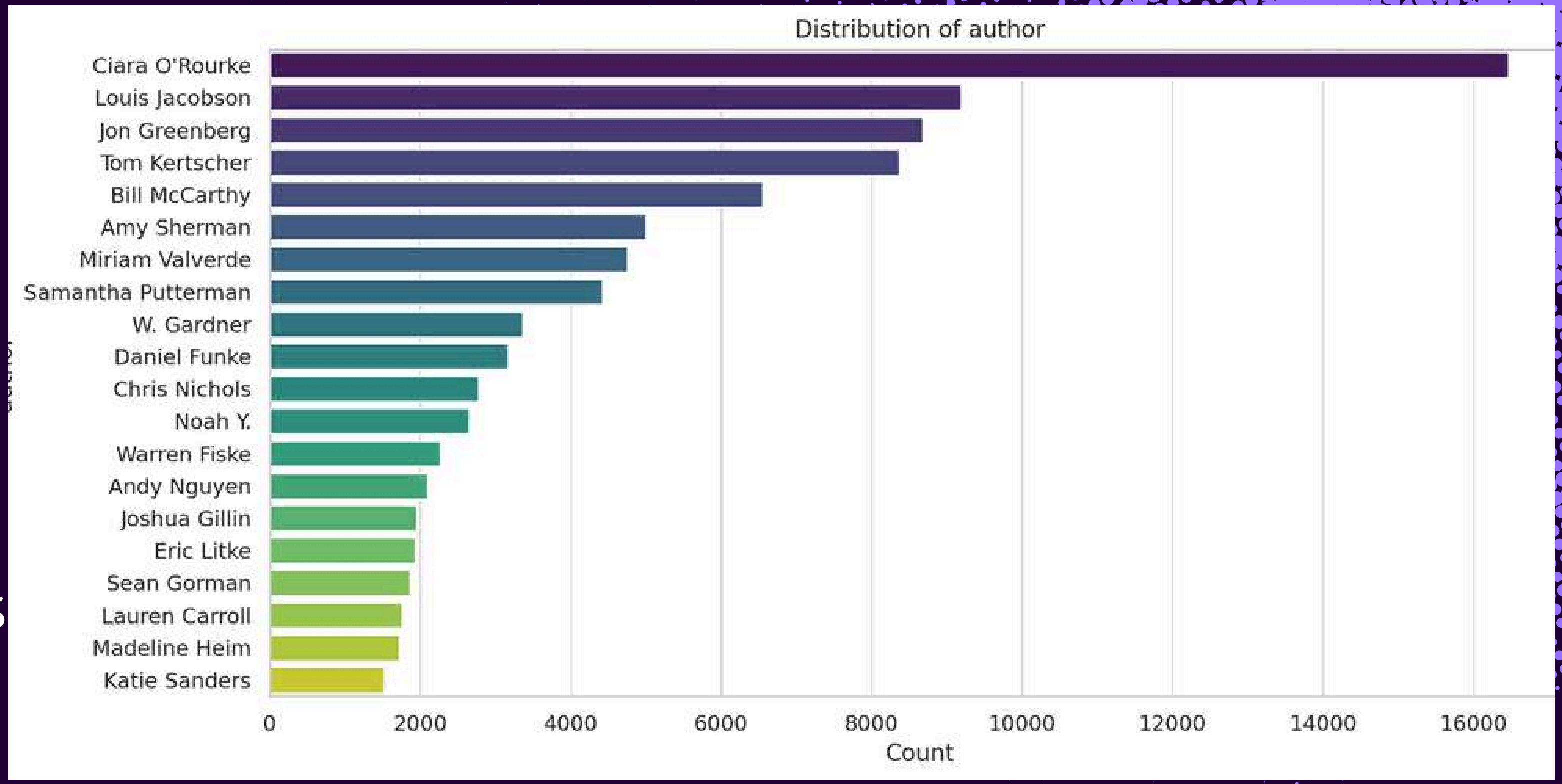
Data Splitting



DATA ANALYSIS: TRUTH SEEKER DATASET

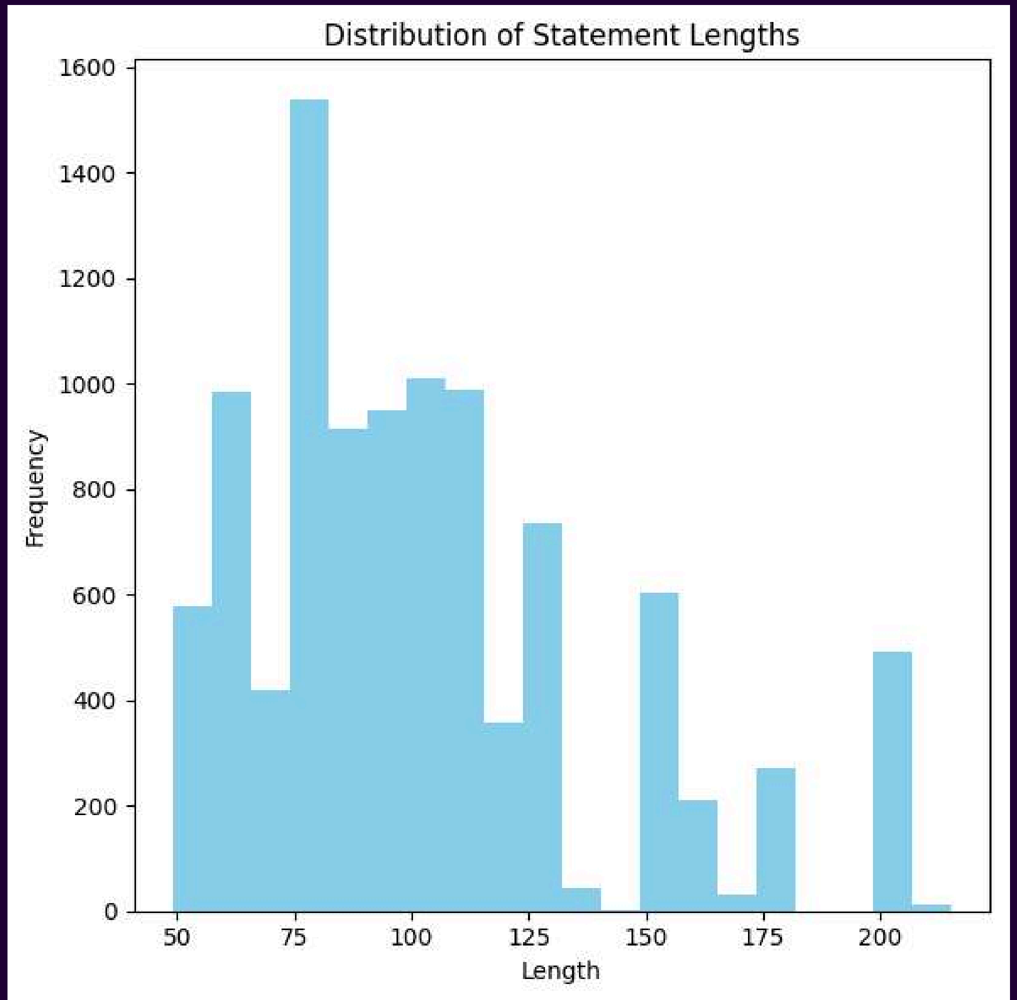


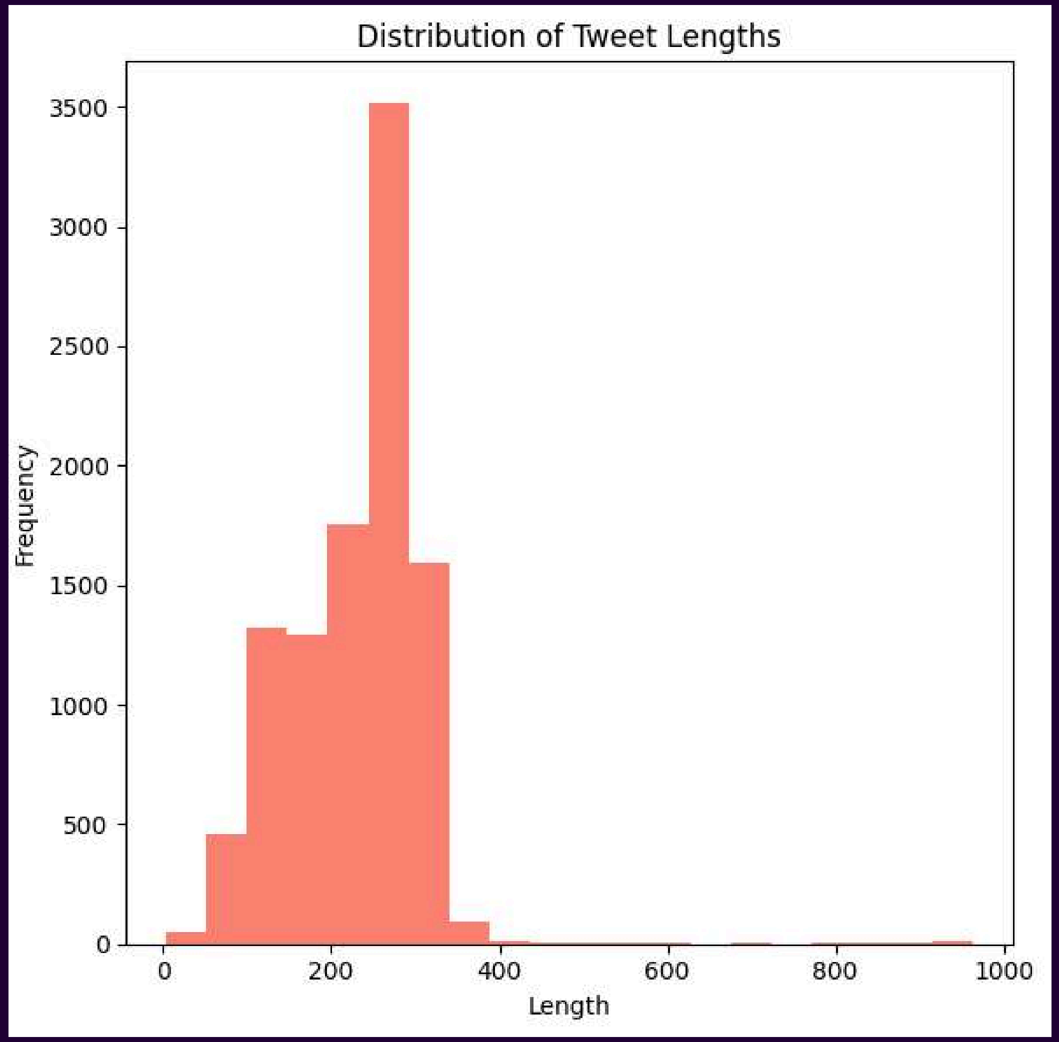
Word Cloud Analysis: Visual representation of frequently used words in fake and real news



Author : Identifies frequent author spreading fake news

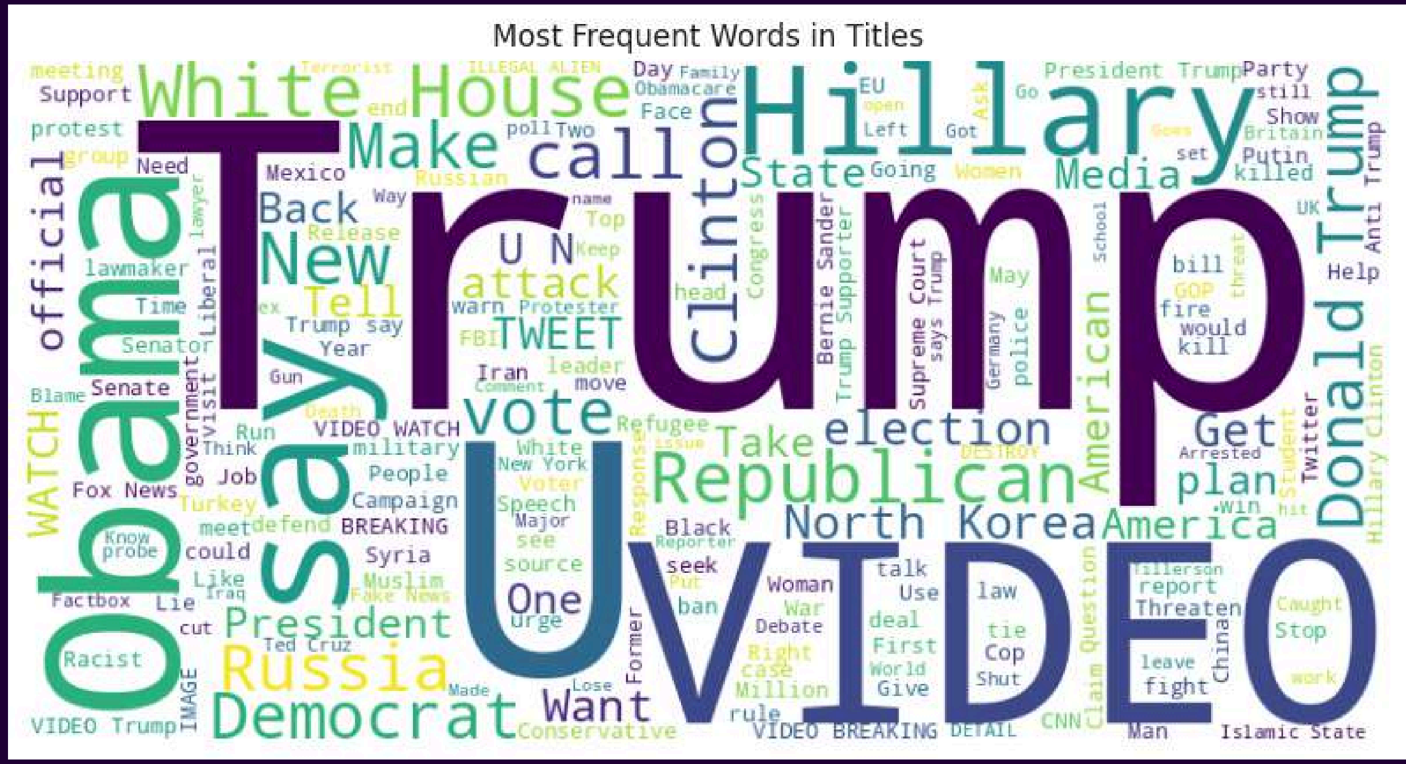
Statement Analysis: Evaluate linguistic patterns and sentiments in news statements



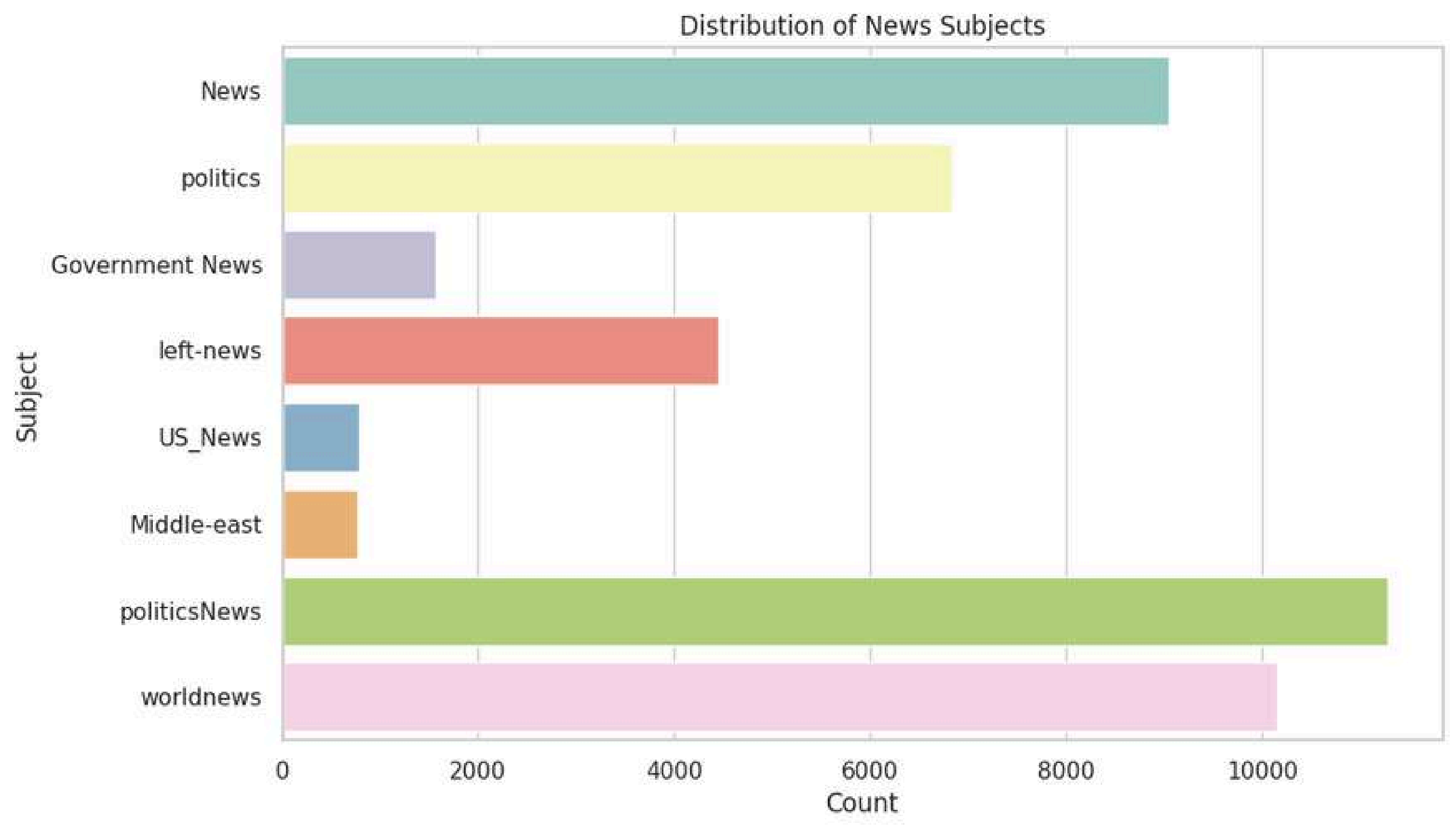


Tweet Analysis: Analysis tweet length, engagement metrics

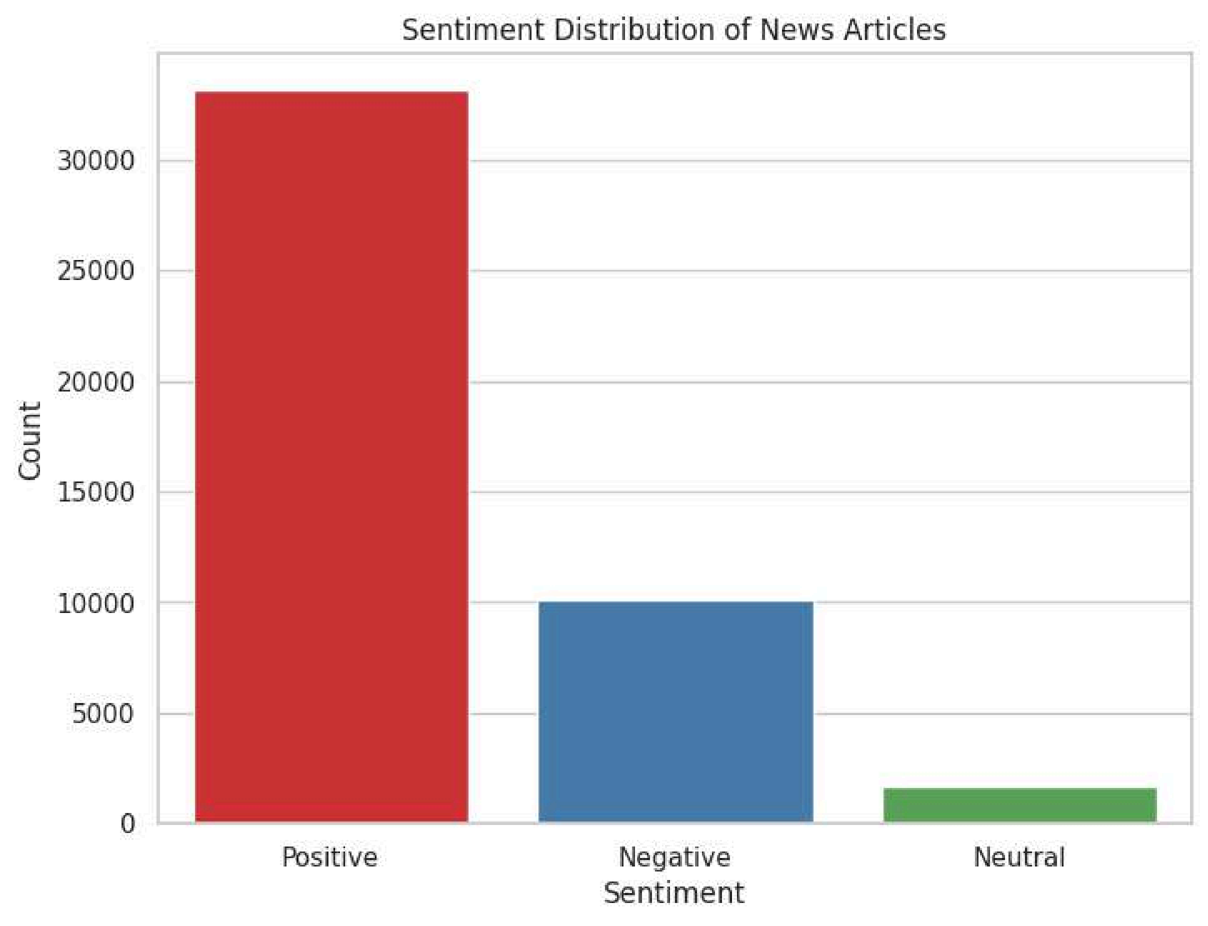
DATA ANALYSIS: ISOT DATASET



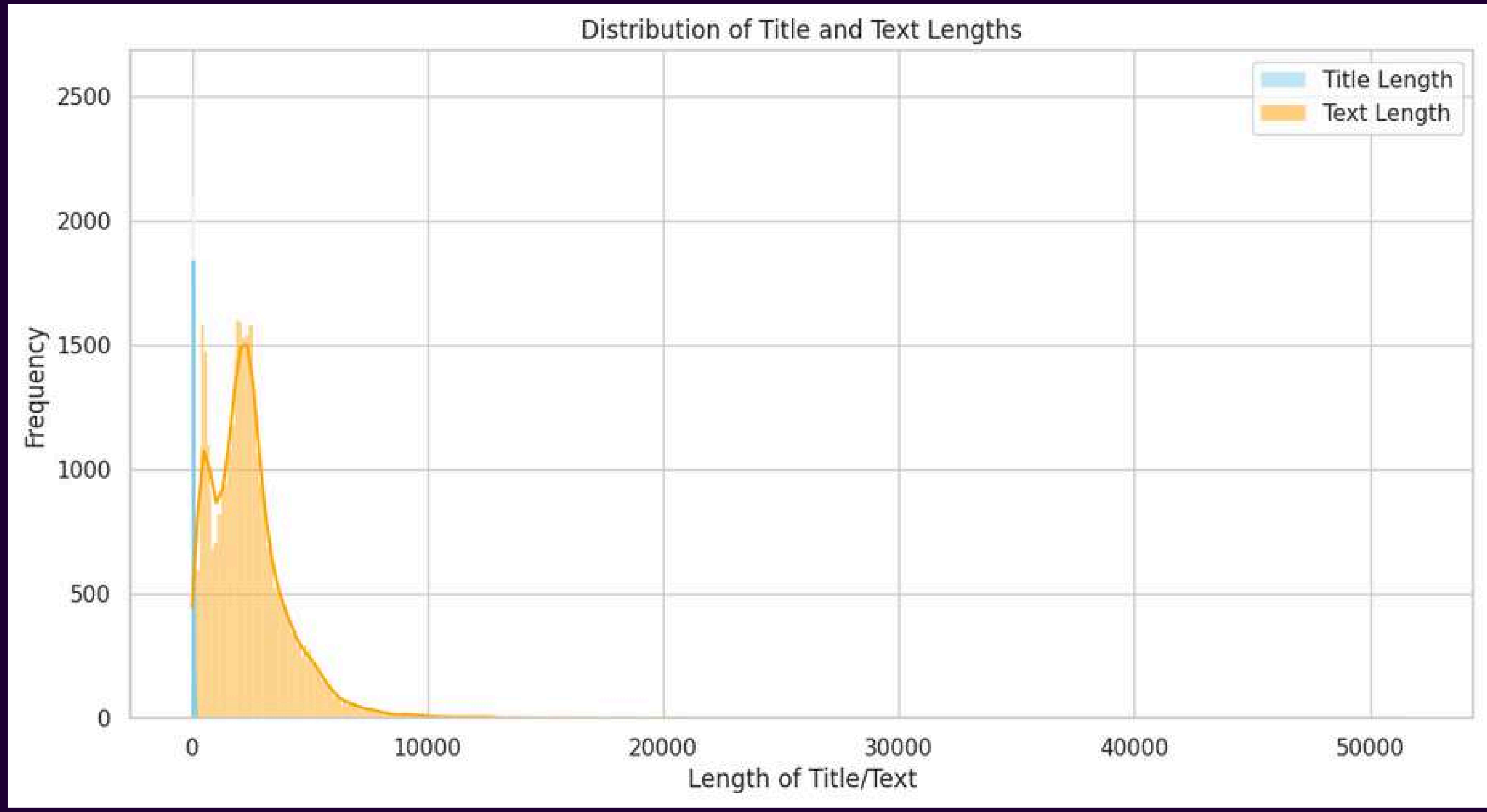
-**Subject Analysis:** Categories news topics (politics, business, entertainment, etc.)



Word Cloud Analysis: Visual representation of frequently used words in fake and real news

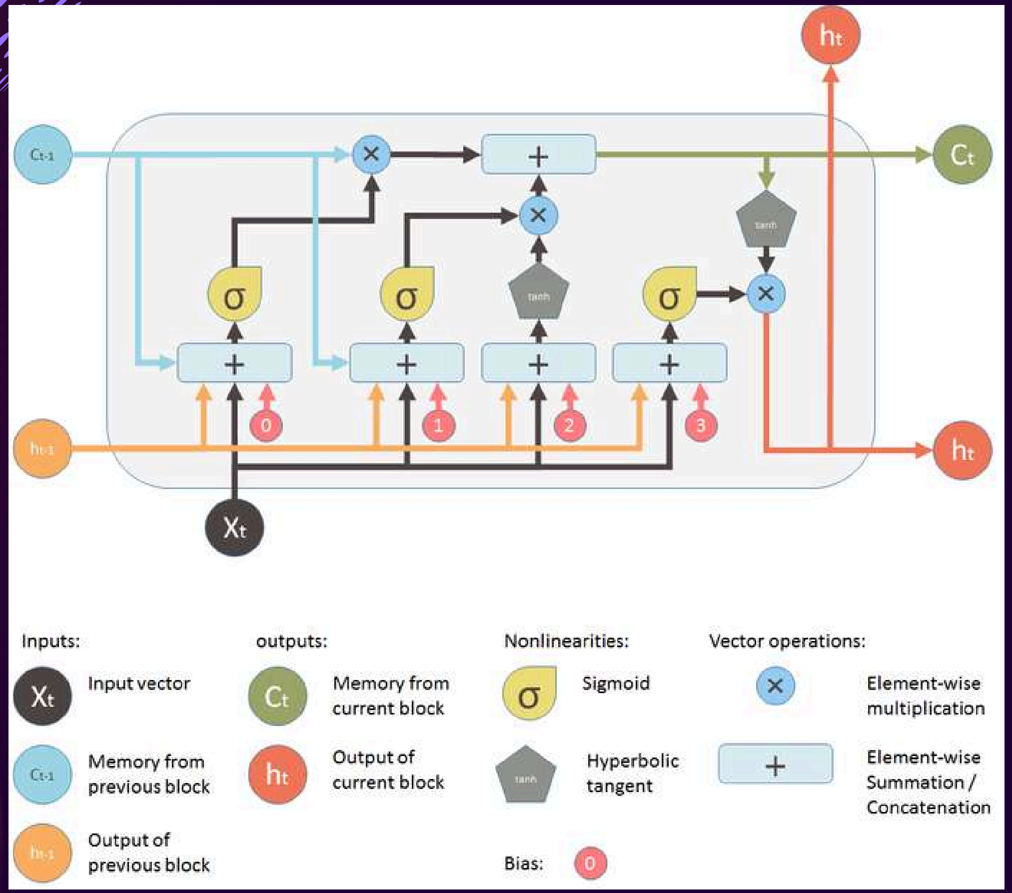


Text Analysis: Analysis the length, complexity and sentiment of the news articles

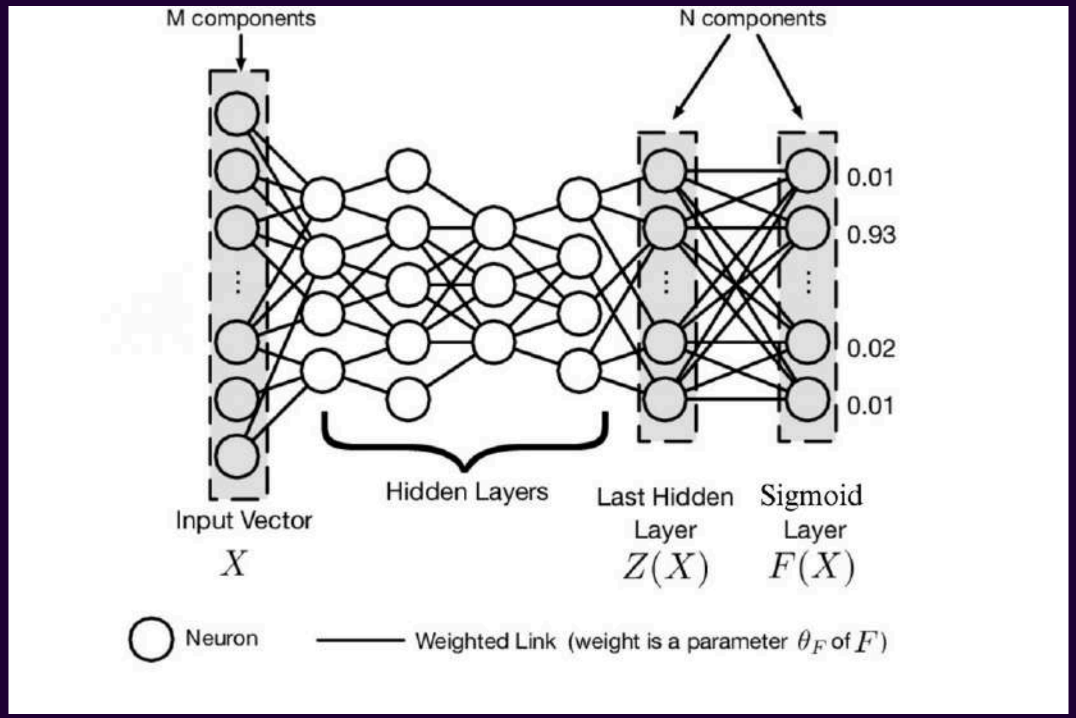


Title Analysis: Examines title structure in fake vs. real news

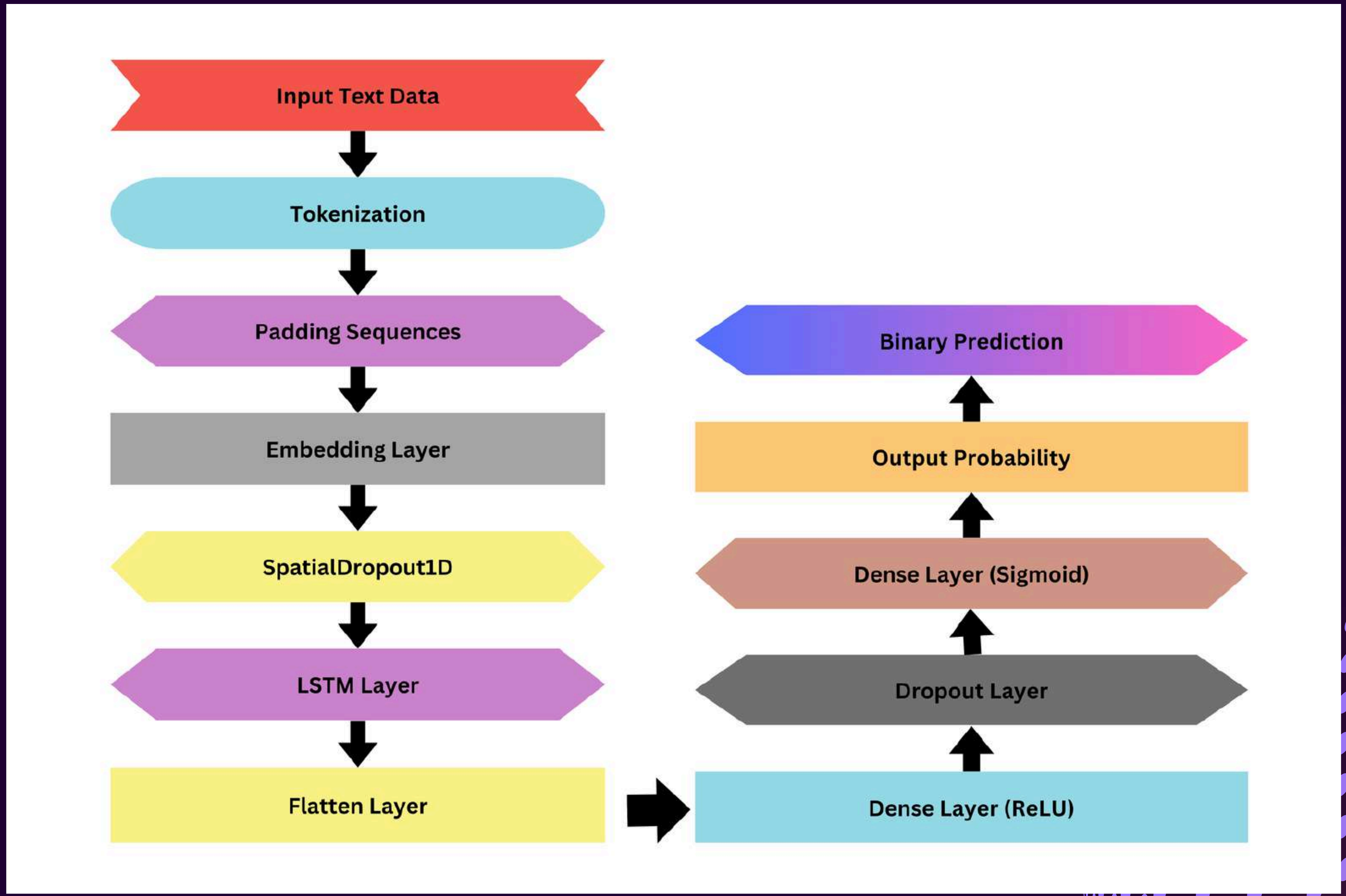
DEEP LEARNING MODELS



LSTM

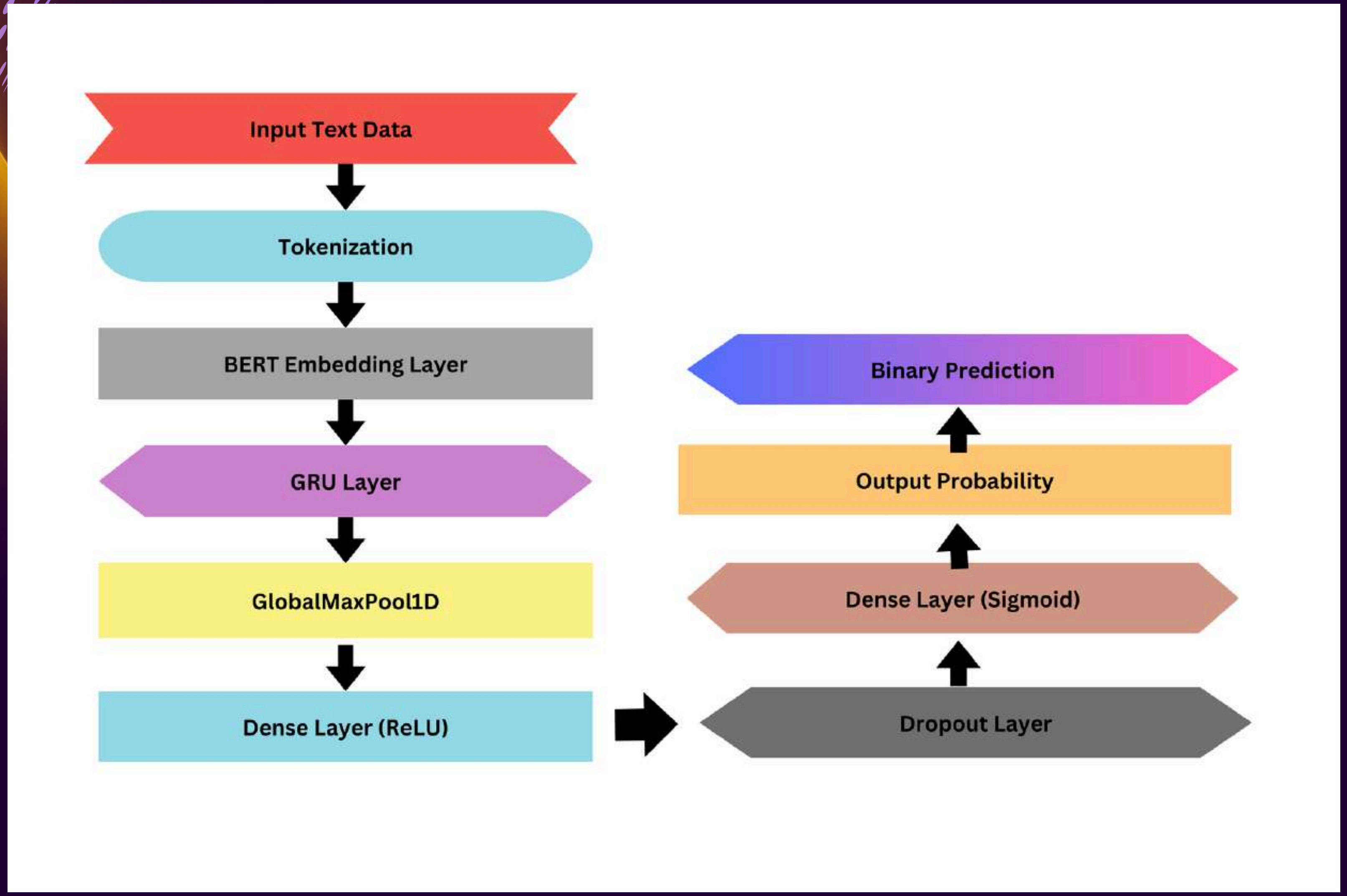


DNN

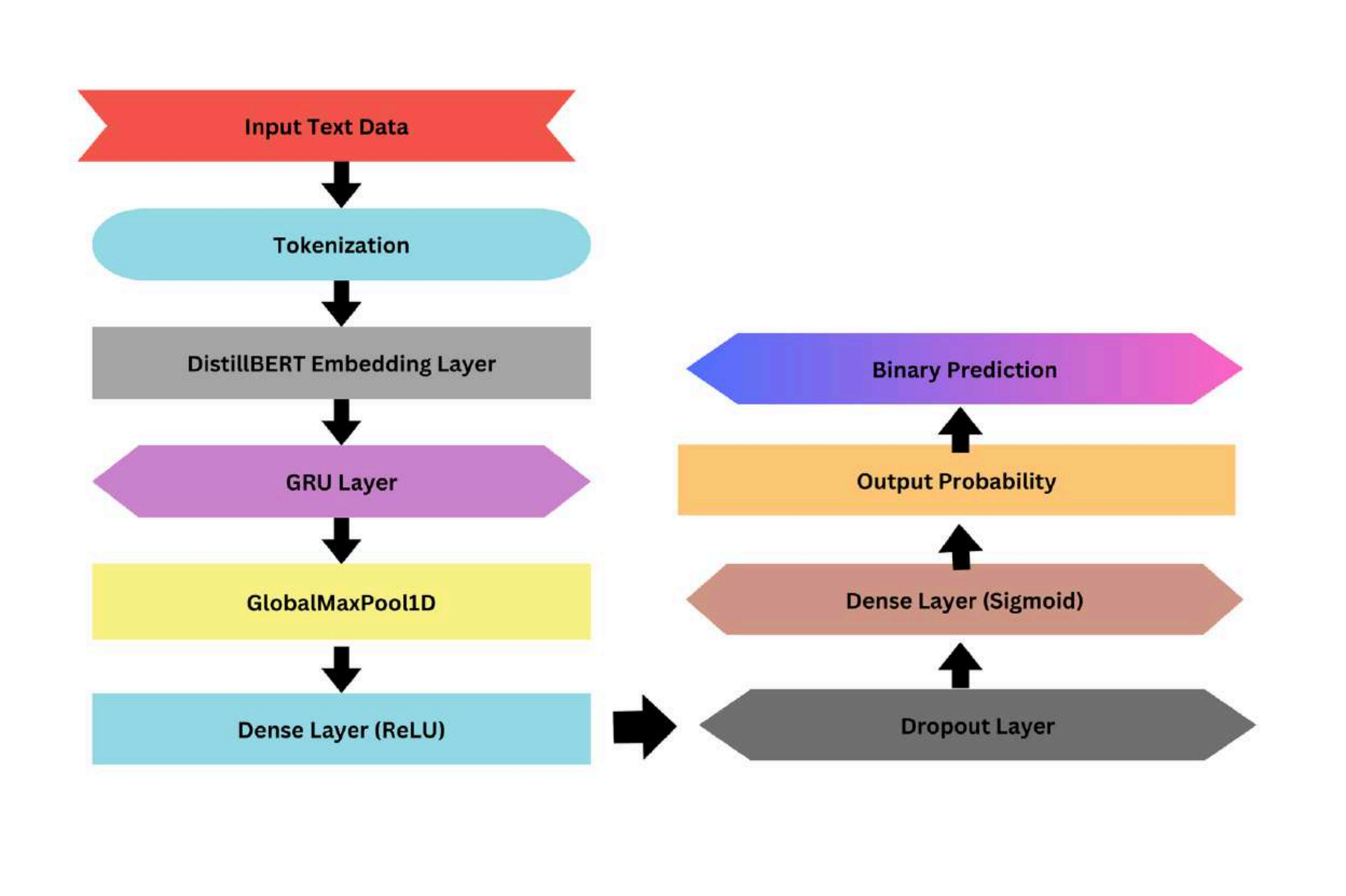


HYBRID LSTM & DNN

ADVANCED DEEP LEARNING MODELS

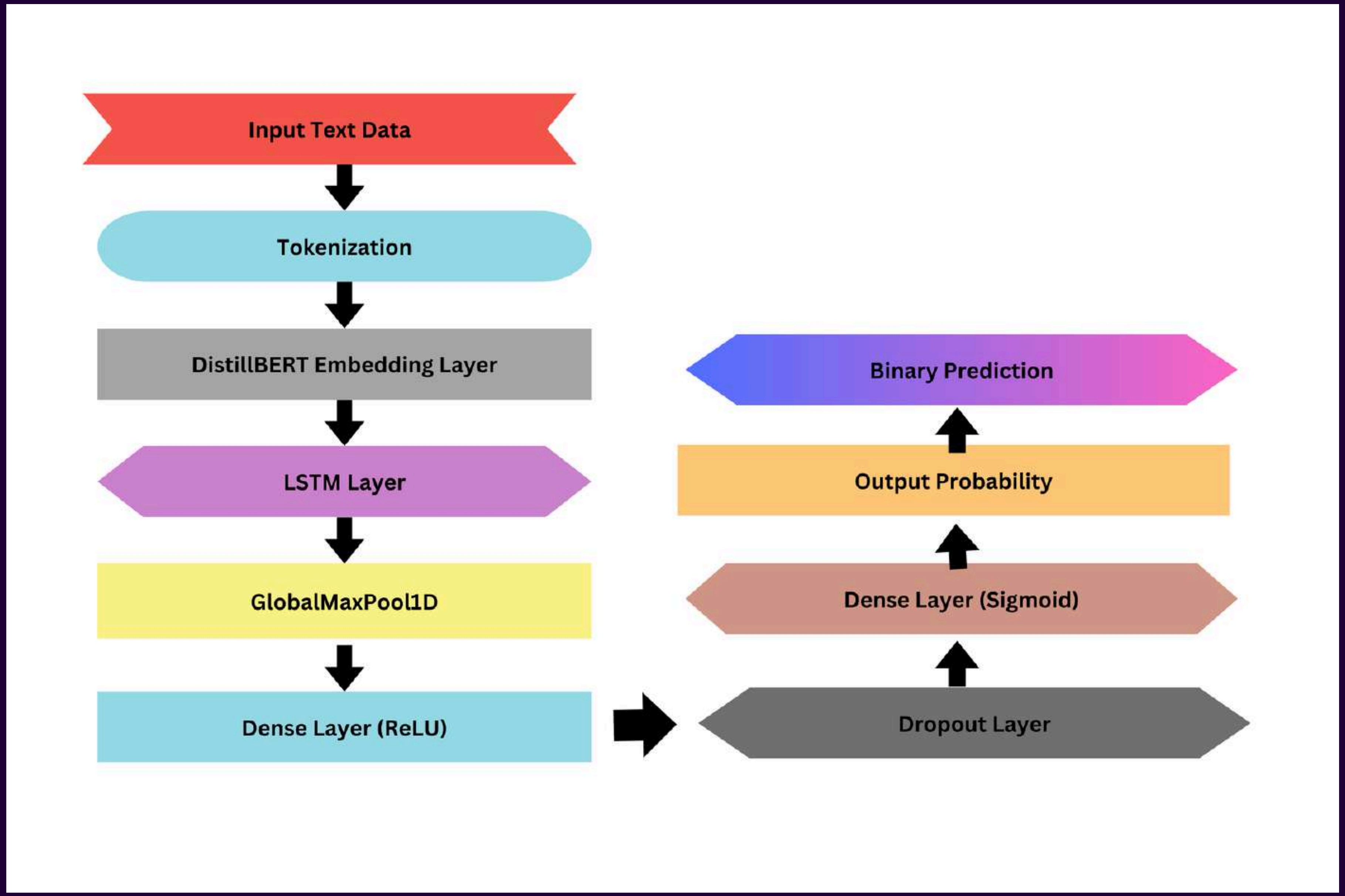


BERT + GRU +DNN



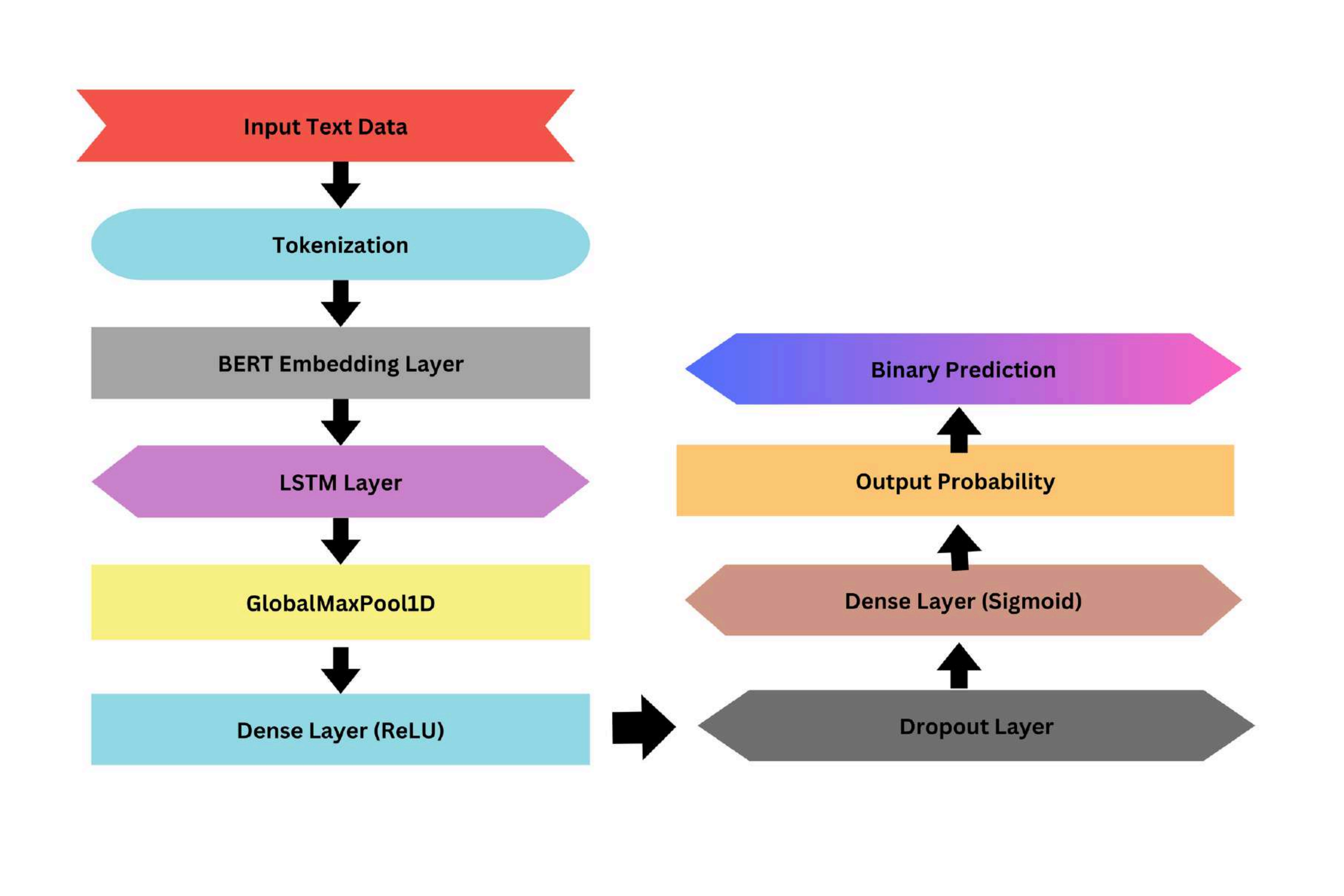
DISTILBERT + GRU +DNN

ADVANCED DEEP LEARNING MODEL



DISTILBERT + LSTM +DNN

PROPOSED TRUTHFORGE MODEL



RESULT EVALUATATION

USED MODELS

- DNN
- LSTM
- Hybrid LSTM & DNN
- BERT +GRU+DNN
- DistilBERT +GRU+DNN
- DistilBERT+LSTM+DNN
- TruthForge

METRICS

- Accuracy
- Precision
- Recall
- F1 score

PERFORMANCE LSTM , DNN

	Precision	Recall	F1-score
Weighted Avg	85.32%	87.22%	86.26%
Accuracy	85.75%		

Table 5.1: Performance Metrics of LSTM Model

	Precision	Recall	F1-score
Weighted Avg	85.23%	86.82%	86.02%
Accuracy	85.52%		

Table 5.2: Performance Metrics od DNN Model

PERFORMANCE HYBRID LSTM & DNN

Max_Length	Epoch	Train_Loss	Val_Loss	Accuracy	Test_Accuracy	Test_F1-score	Test_Precision	Test_Recall
20	10	0.2722	0.2546	88.13%	89.66%	90.26%	87.08%	93.68%
25	10	0.2372	0.2214	91.15%	91.62%	92.06%	89.31%	94.99%
50	10	0.1241	0.1036	95.38%	96.71%	96.78%	96.91%	96.65%

Table 5.3: Performance Metrics of Hybrid LSTM+DNN Model for Different Max Lengths

Max_Length	Epoch	Train_Loss	Val_Loss	Accuracy	Test_Accuracy	Test_F1-score	Test_Precision	Test_Recall
20	10	0.6296	0.5211	71.37%	93.32%	92.95%	93.38%	92.52%
25	10	0.7025	0.6206	70.55%	90.93%	90.29%	91.96%	88.68%
50	10	0.8097	0.7093	58.30%	71.36%	73.68%	65.44%	84.30%

Table 5.4: Performance Metrics of Hybrid LSTM+DNN Model for Different Max Lengths

PERFORMANCE BERT+GRU+DNN

Max_ Length	Epoch	Train_ Loss	Val_Loss	Accuracy	Test_ Accuracy	Test_ F1-score	Test_ Precision	Test_ Recall
20	8	0.3369	0.2802	85.43%	88.73%	88.84%	90.11%	87.60%
24	8	0.3031	0.2390	87.71%	90.06%	89.85%	91.57%	86.02%
28	8	0.2587	0.2110	89.47%	92.44%	92.62%	92.49%	92.75%

Table 5.5: Performance Metrics of BERT+GRU+DNN Model for Different Max Lengths

Max_ Length	Epoch	Train_ Loss	Val_Loss	Accuracy	Test_ Accuracy	Test_ F1-score	Test_ Precision	Test_ Recall
20	8	0.3323	0.2499	86.24%	95.96%	96.76%	95.48%	96.04%
24	8	0.3320	0.2382	87.01%	97.99%	97.89%	97.55%	98.24%
28	8	0.3211	0.2579	86.67%	97.07%	96.97%	95.17%	98.75%

Table 5.6: Performance Metrics of BERT+GRU+DNN Model for Different Max Lengths

PERFORMANCE DISTILBERT + GRU + DNN

Max_Length	Epoch	Train_Loss	Val_Loss	Accuracy	Test_Accuracy	Test_F1-score	Test_Precision	Test_Recall
20	8	0.3004	0.2669	86.94%	88.96%	88.81%	92.25%	85.61%
24	8	0.2690	0.2299	89.00%	90.68%	90.48%	94.75%	86.58%
28	8	0.2442	0.2014	90.38%	92.40%	92.81%	89.97%	95.82%

Table 5.7: Performance Metrics of DistilBERT+GRU+DNN Model for Different Max Lengths

Max_Length	Epoch	Train_Loss	Val_Loss	Accuracy	Test_Accuracy	Test_F1-score	Test_Precision	Test_Recall
20	8	0.2971	0.2273	88.27%	96.12%	95.92%	95.99%	95.85%
24	8	0.3464	0.2447	89.31%	97.40%	97.32%	97.35%	97.29%
28	8	0.3578	0.2832	87.54%	97.34%	97.26%	96.72%	97.81%

Table 5.8: Performance Metrics of DistilBERT+GRU+DNN Model for Different Max Lengths

PERFORMANCE DISTILBERT + LSTM + DNN

Max_ Length	Epoch	Train_ Loss	Val_Loss	Accuracy	Test_ Accuracy	Test_ F1-score	Test_ Precision	Test_ Recall
20	8	0.2885	0.2583	87.77%	89.53%	89.87%	89.17%	90.59%
24	8	0.2555	0.2199	90.17%	90.88%	90.82%	93.84%	87.99%
28	8	0.2210	0.1931	91.23%	93.04%	93.34%	91.30%	95.47%

Table 5.9: Performance metrics of DistilBERT+LSTM+DNN Model for Different Max Lengths

Max_ Length	Epoch	Train_ Loss	Val_Loss	Accuracy	Test_ Accuracy	Test_ F1-score	Test_ Precision	Test_ Recall
20	8	0.2891	0.2238	88.16%	96.22%	96.04%	95.76%	96.32%
24	8	0.3448	0.2758	88.30%	97.27%	97.18%	97.33%	97.03%
28	8	0.3555	0.2712	89.91%	97.63%	97.56%	97.41%	97.71%

Table 5.10: Performance Metrics of DistilBERT+LSTM+DNN Model for Different Max Lengths

PERFORMANCE TRUTHFORGE

Max_ Length	Epoch	Train_ Loss	Val_Loss	Accuracy	Test_ Accu- racy	Test_ F1- score	Test_ Preci- sion	Test_ Recall
20	8	0.3131	0.2707	86.51%	88.90%	88.89%	91.06%	86.83%
24	8	0.2714	0.2325	88.45%	91.05%	91.23%	91.57%	90.87%
28	8	0.2426	0.2015	90.22%	92.81%	93.01%	92.50%	93.53%

Table 5.11: Performance Metrics of TruthForge Model for Different Max Lengths

Max_ Length	Epoch	Train_ Loss	Val_Loss	Accuracy	Test_ Accu- racy	Test_ F1- score	Test_ Preci- sion	Test_ Recall
20	8	0.3204	0.2446	86.45%	96.06%	95.88%	95.45%	96.32%
24	8	0.2326	0.1418	91.28%	98.07%	97.97%	97.82%	98.12%
28	8	0.3124	0.2258	87.06%	98.20%	98.13%	97.46%	98.81%

Table 5.12: Performance Metrics of TruthForge Model for Different Max Lengths

BEST MODEL

Models	Truth Seeker				ISOT			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
LSTM+ DNN	95.38%	96.91%	96.65%	96.78%	71.36%	65.44%	84.30%	73.68%
Bert+ GRU+ DNN	92.44%	92.49%	92.75%	92.62%	97.07%	95.17%	98.75%	96.97%
Distil BERT+ GRU+ DNN	92.40%	89.97%	95.82%	92.81%	97.34%	96.72%	97.81%	97.26%
Distil BERT+ LSTM+ DNN	93.04%	91.30%	95.47%	93.34%	97.63%	97.41%	97.71%	97.56%
Truth-Forge	92.81%	92.50%	93.53%	93.02%	98.20%	97.46%	98.81%	98.13%

Table 5.13: Comparison between Performance of different Models on Truth Seeker and ISOT datasets

CONCLUSION AND KEY TAKEAWAYS

- Social media spreads both real and fake news, requiring fake news detection
- Deep learning models, hybrid and advanced deep learning models' results were compared .
- TruthForge model is the most efficient
- Achieving 98.2% accuracy

FUTURE WORK

01 Integrate multimodal data (text, images, and videos).

02 Deploy real-time fake news detection system.

03 Improve robustness against adversarial manipulation.

THANK YOU

REFERENCES

H. AHMED, I. TRAORE, AND S. SAAD, "DETECTING OPINION SPAMS AND FAKE NEWS USING TEXT CLASSIFICATION," SECURITY AND PRIVACY, VOL. 1, E9, DEC. 2017. DOI: 10.1002/SPY2.9.

N. DELIGIANNIS, T. HUU, D. M. NGUYEN, AND X. LUO, "DEEP LEARNING FOR GEOLOCATING SOCIAL MEDIA USERS AND DETECTING FAKE NEWS," IN NATO WORKSHOP, 2018.

K. POPAT, S. MUKHERJEE, A. YATES, AND G. WEIKUM, "DECLARE: DEBUNKING FAKE NEWS AND FALSE CLAIMS USING EVIDENCE-AWARE DEEP LEARNING," ARXIV PREPRINT ARXIV:1809.06416, 2018.

C. K. HIRAMATH AND G. DESHPANDE, "FAKE NEWS DETECTION USING DEEP LEARNING TECHNIQUES," IN 2019 1ST INTERNATIONAL CONFERENCE ON ADVANCES IN INFORMATION TECHNOLOGY (ICAIT), IEEE, 2019, PP. 411–415.

M. M. SADR ET AL., "THE USE OF LSTM NEURAL NETWORK TO DETECT FAKE NEWS ON PERSIAN TWITTER," TURKISH JOURNAL OF COMPUTER AND MATHEMATICS EDUCATION (TURCOMAT), VOL. 12, NO. 11, PP. 6658–6668, 2021.

S. DADKHAH, X. ZHANG, A. G. WEISMANN, A. FIROUZI, AND A. A. GHORBANI, "THE LARGEST SOCIAL MEDIA GROUND-TRUTH DATASET FOR REAL/FAKE CONTENT: TRUTHSEEKER," IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS, PP. 1–15, 2023. DOI: 10.1109/TCSS.2023.3322303.