Automated Detection of Fake News in Urdu Language Using Pre-Trained Transformer Models

MSc Research Project
Data Analytics

Jackay Lohano
Student ID: 23212896

School of Computing
National College of Ireland

Supervisor: Hamilton V. Niculescu

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Jackay Lohano		
Student ID:	x23212896		
Programme:	Data Analytics	**Year:**	2024
Module:	MSc Research Project		
Supervisor:	Hamilton V. Niculescu		
Submission Due Date:	12-12-2024		
Project Title:	Automated Detection of Fake News in Urdu Language Using Pre-Trained Transformer Models		
Word Count:	7411	**Page Count**	22

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Jackay Lohano

Date: 12-12-2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Automated Detection of Fake News in Urdu Language Using Pre-Trained Transformer Models

Jackay Lohano
x23212896

Abstract

The propagation of misinformation across different languages and domains on various social media platforms is of grave concern for societies and individuals due to its wide-range consequences. Although, researchers have addressed this challenge using advanced deep learning (DL) models, fake news detection in low resource languages such as Urdu is still at nascent stage. Studies have used traditional machine learning (ML) models on a very small and domain-restricted Urdu datasets for fake news detection. This study explores fake news classification in Urdu using three state-of-the-art (SOTA) pre-trained multilingual transformer models i.e. mBERT, DistilmBERT and mT5 on a large and multi-domain Urdu dataset. Models are evaluated on the evaluation metrics such as accuracy, precision, recall and f1-score. The results show that DistilmBERT model demonstrates promising results with an accuracy of 89% compared to its larger counterpart mBERT and mT5 models. The findings reveal the potential of DistilmBERT model for real-world applications in memory-constrained environments for identification of fake news. This research contributes to the ongoing efforts to combat misinformation in resource-scarce languages by building a reliable Urdu fake news detection model, addressing the complexities of information dissemination and manipulation in modern age.

Keywords: *Fake News Detection, Urdu Language, Transformer models, Low resource language, Transfer Learning*

1 Introduction

In today's digital world, social media has gained immense popularity among people for easy access to news and information on various topics. However, it has also raised concerns for the authenticity of the information shared and spread online (Shu et al., 2017). The dissemination of fake news is a pressing problem considering the rate at which information is shared and spread in the digital landscape and often gets more attention than real news. Moreover, fake news can change public opinions, damage a company's reputation, and even influence election results (Grinberg et al., 2019). Among many languages, Urdu spoken by more than 230 million[1] people worldwide, is faced with unique challenges in combating misinformation due to linguistic complexities and the lack of annotated datasets. Despite the growing prevalence of fake news in Urdu, existing detection systems are predominantly tailored to high-resource languages like English. This gap hinders the effectiveness of automated detection systems in accurately detecting fake news within Urdu content (Amjad et al., 2020). A robust automatic

[1] https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide/

system is required that can understand the intricate linguistic nuances of Urdu text for effective fake news detection since the manual detection process demands more time, resources, and cost.

While researchers have used several ML and DL approaches for fake news classification in Urdu, they lack in several aspects. The datasets used for fake news detection in Urdu contain a limited number of samples and do not incorporate multiple domains. Amjad et al. (2020) detected fake news in Urdu using a dataset containing only 900 news items from five domains. A similar work was done by Akhter et al. (2021) with 2000 news items. The limited sample size and domain-restricted news data are not suitable for real-world applications, as in real life news can spread from a variety of domains. Moreover, the techniques applied to identify fake news in Urdu mainly rely on traditional ML approaches (Akhter et al., 2021). Therefore, considering the existing work and the need to fill the research gap, this study aims to investigate fake news detection in Urdu with a large dataset covering news from a variety of domains using advanced pretrained language models based on transformer architecture. For this purpose, this study uses a recently published benchmark Urdu dataset, "Ax-to-Grind" for Urdu fake news detection, providing a promising foundation for advancing research in this field (Harris et al., 2023).

Research Question: *How accurately can multilingual versions of pre-trained transformer-based models detect the fake news in a low-resource language Urdu?*

This study aims to bridge the gap by leveraging Transfer Learning (TL) on three transformer models for Urdu fake news classification and is an attempt to further the research in low-resource languages by building reliable fake news detection system. This study can benefit policymakers, news media organizations, content moderators, and other relevant stakeholders in their efforts to stop the proliferation of fake news, thereby fostering a more informed society. Specifically, the objectives of this study are to:

- Prepare Ax-to-Grind Urdu dataset for modelling and perform text analysis to understand various features of news data helpful at implementation stage.

- Implement three pre-trained multilingual transformer models (mBERT, DistilmBERT and mT5) for classifying fake and real news using the Ax-to-Grind Urdu dataset.

- Assess and compare the performance of these models based on evaluation metrics such as accuracy, precision, recall, and F1-score.

The rest of the paper is organized as follows: Section 2 presents a review of existing literature work conducted on fake news detection in Urdu and transformer-based approaches for addressing fake news classification. Section 3 outlines the research methodology adopted for this work. Section 4 provides details on the framework and architecture of the models used in this research. Section 5 elaborates on the workflow of the project and its implementation, from data preprocessing to the evaluation process. Section 6 discusses the results obtained from different implementations. Finally, Section 7 concludes this research and discusses future work to extend the current research in the domain of fake news detection in low-resource languages.

2 Related Work

Fake news detection is a popular topic in Natural Language Processing (NLP) tasks and is addressed by many researchers. While there has been more research in high resource languages such as English, researchers are now investigating fake news classification in low resource languages. This section discusses the previous research work conducted to detect fake news in Urdu and advanced techniques based on transformers applied to fake news classification.

2.1 Fake News Detection in Urdu Language

Amjad et al. (2020) contributed the Bend the Truth (BET) benchmark dataset for fake news detection in Urdu, comprising of 900 news articles from five domains. The authors implemented ML models with a variety of feature extraction techniques, such as word n-grams and character n-grams. Due to the limited sample size and domain, the model may not be reliable for fake news detection, as in reality fake news comes from a variety of domains. Another limitation is that fake versions of true news are created intentionally by professional journalists, which is not a good practice for real-world applications. In a further study conducted by Amjad, Sidorov and Zhila (2020), the augmentation technique was applied on BET dataset. The authors added 400 news items to the original dataset by translating English news to Urdu. This study was aimed at investigating the role of augmentation in improving the accuracy of Urdu fake news detection. However, as per the authors, these methods of augmentation and machine translation did not provide any improvement but even decreased the accuracy on the original dataset.

The BET dataset is used by many studies. For example, Balaji and Bharathi (2020) utilized feature extraction techniques such as Term Frequency-Inverse Document Frequency (TF-IDF) and FastText on BET dataset and achieved 78.7% accuracy on test data with Random Forest (RF) classifier, whereas Rafique et al. (2022) achieved a better accuracy of 92% with Bag of Words (BoW) feature technique. Akhter et al. (2021) used ensemble machine learning models for identification of fake news detection in Urdu language. The authors experimented with BET dataset and contributed new dataset comprising of 2000 news items however domain information for new dataset is not provided. Moreover, this new dataset was originally in English which was translated into Urdu using Google Translate and no manual verification was conducted to check accuracy of translated news. The authors implemented three traditional ML models Naive Bayes (NB), Decision Tree (DT) and Support Vector Machines (SVM) and five ensemble models however the results showed a lower accuracy and could not build an accurate detection model (Akhter et al., 2021). Farooq et al. (2023) established a new dataset consisting of 4,097 news items in Urdu language from nine domains. This study proposed an ensemble classifier Extra Tree (ET) and RF as base learner and Logistic Regression (LR) as final learner and achieved 93% accuracy. However, the fake news dataset was collected from Urdu translations of English fake news data and was manually verified.

Some researchers have used deep learning methods for feature extraction from Urdu text. Lin et al. (2020) implemented Character-level Convolutional Neural Networks (CharCNN) for character level feature extraction and Robustly Optimized BERT Approach (RoBERTa) for word level feature extraction and evaluated the model on BET Urdu dataset with an accuracy of 90%. On the other hand, Reddy et al. (2020) used Bidirectional Gated Recurrent Unit (Bi-GRU) for feature representations and obtained test accuracy of 81.7% and F1-score of 80.7%. Kishwar and Zafar (2022) investigated fake news detection in Pakistan and developed a benchmark dataset consisting of 11,000 news items. The authors achieved F1-score of 0.94, however, the news is not in Urdu language but English. This study demonstrates that large datasets outperform limited sized dataset for fake news detection. A similar work is conducted by Kareem and Awan (2019) on Pakistani news however the news is in English language.

2.2 Transformer-Based Approaches for Fake News Detection

Recently, researchers have focused more on using pre-trained language models for fake news classification due to deeper contextual understanding of input text gained from attention mechanism. Ranjan and Agrawal (2022) used transformer models like BERT, XLNet, RoBERTa and Longformer with Genetic Algorithm and Information Gain for identification of fake news using LIAR dataset which contained 12,386 English news items from various sources. Qazi et al. (2020) compared the performance of Hybrid-CNN model with transformer model for fake news detection on LIAR dataset. They found that transformer model improved the accuracy by 15% as compared to Hybrid-CNN model. Alghamdi et al. (2023) classified fake news into six categories such as mostly true and mostly fake because often the whole news is not fake but some part of it might be tailored. They implemented BERT model and evaluated its performance on LIAR dataset. In another study by Jwa et al. (2019) authors adopted BERT model for fake news detection by analysing the relationship between title and content of the news item and achieved a F1-score of 74%. Baruah et al. (2020) proposed BERT model for classifying fake tweets and real tweets which were in English language and obtained an accuracy of 69% on test set.

The language understanding capabilities of transformer architecture and their extensive use to address fake news detection task in resource rich languages has developed interest of researchers to use transformer models on resource-scarce languages. For example, Azizah et al. (2023) developed fake news detection system in Bahasa which is the official language of Indonesia using transformer models like multilingual BERT, ALBERT, IndoBERT and RoBERTa. Out of which, ALBERT showed better results with an accuracy of 87% as compared to other models. In another studyby Rahman et al. (2022), fake news detection in Bengali language employing various ML, DL and transformer-based approaches is investigated. The findings suggested that Multilingual Masked Language Model (XLM-R) outperformed all other models, achieving F1-score of 98%. De et al. (2022) proposed multilingual BERT model for fake news detection in five languages and seven domains. Authors conducted variety of experiments with domain-specific and language-specific and found high accuracy with domain-specific and domain agnostic experiments. Panda and Levitan (2021) compared performance of four transformer models to address COVID-19 misinformation in three languages English, Arabic and Bulgarian. Authors found that English BERT model performed best for English language and its multilingual variant mBERT showed best results for Arabic and Bulgarian. These studies reveal how well the multilingual variant of BERT model (mBERT) can generalize across other languages with limited linguistic resources without compromising on performance. In a recent study conducted by Harris et al. (2023), an ensemble technique was explored using three transformer models mBERT, XLNet and XLM-RoBERTa for fake news detection in Urdu. The model was evaluated with large multi-domain dataset, achieving 95% accuracy. This study shows the effectiveness of transformer model for fake news detection in low-resource languages.

2.2 Summary

From the above literature analysis, it can be concluded that fake news classification in Urdu language faces many limitations. First limitation is the dataset used for Urdu fake news detection. Most studies have used BET dataset while others have modified it to include more content by translating English news to Urdu or creating fake version of real news. In all cases, the dataset is limited in number of samples and contains news from fewer domains. Second limitation is the implementation techniques. Most studies discussed in above literature have used traditional ML and DL approaches to detect fake news in Urdu language. Despite good results on small datasets, these models are not able to capture linguistic features of text data. Although, some researchers have implemented transformer models for fake news detection on multilingual datasets of resource-scarce languages, very little work is conducted for fake news detection in Urdu language with advanced language models. Transformer architecture contains self-attention mechanism which provides rich understanding of context of input text. This study takes this advantage and fills gap by implementing three pre-trained language models based on transformer architecture and evaluates their performance on Ax-to-Grind Urdu dataset which is first publicly available large dataset with multi-domain news items. The comparative analysis of literature studies for identification of fake news in Urdu language along with the implementation techniques, model accuracy and limitations are given in Table 1. This also lists the contribution of this research by proposing implementation of mBERT, DistilmBERT and mT5 models in the field of fake news detection in Urdu language.

Reference	Technique	Accuracy	Dataset	Domains	Limitation
Amjad et al. (2020)	LR, SVM, DT, AdaBoost	88%	900	5	Limited sample size and domains
Amjad, Sidorov and Zhila (2020)	SVM, AdaBoost	65% 75%	400 1300	5	Machine Translated (MT) dataset. Combined original BET dataset with MT dataset No manual verification. Domain-restricted.
Akhter et al. (2021)	NB, DT, SVM	83.3%	2000	5	Machine translated dataset. No manual verification.
Lin et al. (2020)	RoBERTa, charCNN	90%	900	5	Limited sample size and domains

Farooq et al. (2023)	KNN, RF, ET, SVM, LR and Stacked Model	93.39%	4097	9	Imbalanced dataset with no technique applied to balance.
Harris et al. (2023)	Ensemble model with mBERT, XLNet, XLM-RoBERTa	95.6%	10,083	15	Requires more computing power due to ensemble of multiple models.
This work	mBERT, DistilmBERT, mT5	-	10,083	15	-

Table 1: Comparative analysis of literature for Urdu fake news detection

3 Research Methodology

This research employs Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology. This methodology is used by researchers to build reliable fake news detection model (Shariff et al., 2022). It involves various stages as shown in Figure 1 which are briefly described.

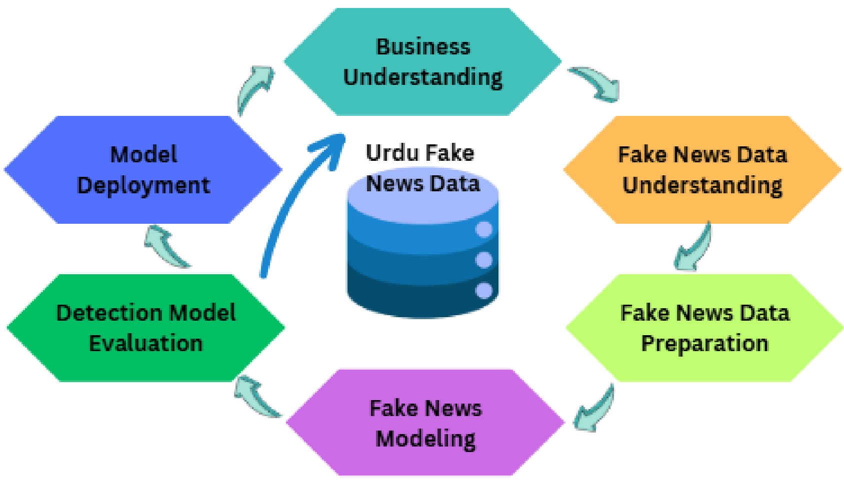

Figure 1: CRISP-DM Methodology for Urdu fake news detection

3.1 Business Understanding

Business understanding is the first stage of CRISP-DM methodology. It focuses on understanding the requirements and goals of research problem. According to Vosoughi et al. (2018), fake news spreads six times faster than real news on social and 70% people are not able to distinguish between real and fabricated content. The effect of fake news propagation is tremendous on societies and individuals as well (Aïmeur et al., 2023). In this research, the objective is to build an automated model for fake news identification in Urdu which can be helpful tool for various stakeholders including government bodies, and media organizations to curb the proliferation of fake information and avoid its consequences.

3.2 Data Understanding and Data Selection

The dataset selected for this research work is taken from a recent paper by Harris et al. (2023). The dataset published by this paper "Ax-to-Grind: Benchmark Dataset for Urdu Fake News Detection" is first largest dataset in Urdu available publicly. It contains 10,083 total news items from 2017 to 2023 across fifteen different domains, giving more realistic approach to detect fake news in Urdu language as in real time fake news spreads from a variety of domains. In addition, this dataset is manually annotated and verified by professional journalists and is sourced from various reliable Urdu newspapers in India and Pakistan.

3.3 Data Preparation and Preprocessing

The raw textual data contains many irrelevant characters and words which are removed to lessen the input space and improve model's performance. It starts from cleaning data such as to remove stop words, special characters, duplicated records and missing entries present in the dataset (Sun et al., 2018). Previous studies have used Urduhack library (Arshad et al., 2022; Tahir and Mehmood, 2022) for Urdu text processing. However, this library does not provide more functions such as stop word removal. Therefore, this study utilizes LughaatNLP[2] library which provides complete toolkit for Urdu text preprocessing.

3.4 Modelling

The cleaned dataset is split into train, validation and test sets. The models are trained on training data. In this study, multilingual versions of three pretrained transformer models are used. Following is the brief description of the models implemented in this study. The architecture and implementation are discussed in detail in later sections.

3.4.1 mBERT and DistilmBERT

Multilingual Bidirectional Encoder Representation from Transformers (mBERT) is the multilingual extension of original BERT model, introduced by Devlin et al. (2019). Unlike BERT, which is trained on English data, mBERT model is trained on Wikipedia pages of 104 languages. BERT uses bidirectional learning meaning that it learns the context of the text from both left and right as opposed to unidirectional model. Bhawal and Roy (2021) used mBERT model but on very limited Urdu dataset comprising of 1300 news items from five domains. Based on training parameters, mBERT has variants such as mBERT Base and mBERT Large. This study investigates the performance of mBERT Base model on a much larger dataset from multiple domains. The detail on the architecture is discussed in section 4.

Distilled version of BERT model (DistilBERT) is the lightweight BERT model introduced by Sanh et al. (2019), which retains 97% language understanding capabilities at a reduced size by 40% and is 60% faster as compared to BERT model. The multilingual variant of English

---

[2] https://github.com/MuhammadNoman76/LughaatNLP

DistilBERT is DistilmBERT which is trained on Wikipedia pages from 104 languages including Urdu. Due to its smaller size, DistilmBERT model is used by many researchers for various tasks. Ullah et al. (2024) employed DistilmBERT model for hate speech and offensive content detection in Telugu language. To the best of author's knowledge in the light of literature, DistilmBERT model has not been applied for Urdu fake news detection. This study implements base variant of DistilmBERT model and compares its performance with mBERT on Urdu dataset.

3.4.2 mT5

Text-to-Text Transfer Transformer (T5) model transforms all language tasks into text-to-text format where input and output both are in text. (Raffel et al., 2020). T5 is trained on English data and its multilingual variant mT5 is trained on 101 languages, including Urdu (Xue, 2020). Sabry et al. (2022) leveraged T5 model for detection of hate speech and offensive content in English tweets. T5 model has many variants based on training parameters such as T5 Small and T5 Large. This study implements mT5 small model and aims to investigate its performance on dataset Ax-to-Grind for fake news detection in Urdu which has not been implemented before to the best of author's knowledge.

3.5 Evaluation Process

The fake news detection is a binary classification problem where the performance of all models is evaluated using evaluation metrics accuracy, precision and recall value and F1-score as evaluated by previous researchers (Farooq et al., 2023).

3.5.1 Accuracy

It gives the proportion of total accurate predictions in the data. It is helpful to assess overall performance of the model. The formula is given by following equation where TP is True Positives, TN is True Negatives, FP is False Positives and FN is False Negatives.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

3.5.2 Precision

Precision is the proportion of correct positive predictions among all predicted positives. It focuses on model's ability to predict true positive.

$$Precision = \frac{TP}{TP + FP}$$

3.5.3 Recall

Recall measures the ability of the model to correctly identify all the actual positive instances, expressed as ratio of true positives to the sum of true positives and false negatives.

$$Recall = \frac{TP}{TP + FN}$$

3.5.4 F1-Score

F1 score is the harmonic mean of precision and recall. This metric seeks a balance between precision and recall.

$$F1\ Score = 2\ x\ \frac{Precision\ x\ Recall}{Precision + Recall}$$

4 Design Specification

This section provides brief information on the architecture of the models used in this research study which underlie the implementation of fake news detection model in Urdu. The implementation is further explained in detail in Section 5.

4.1 Transformer Architecture

The transformer model was first introduced in the paper "Attention Is All You Need" authored by Vaswani et al. (2017). It follows encoder-decoder configuration entirely based on attention mechanism instead of recurrence and convolutions used in other models such as RNN and LSTM. Attention phenomenon in transformers was an innovation as now models can focus on different parts of the input data giving them better understanding of the context.

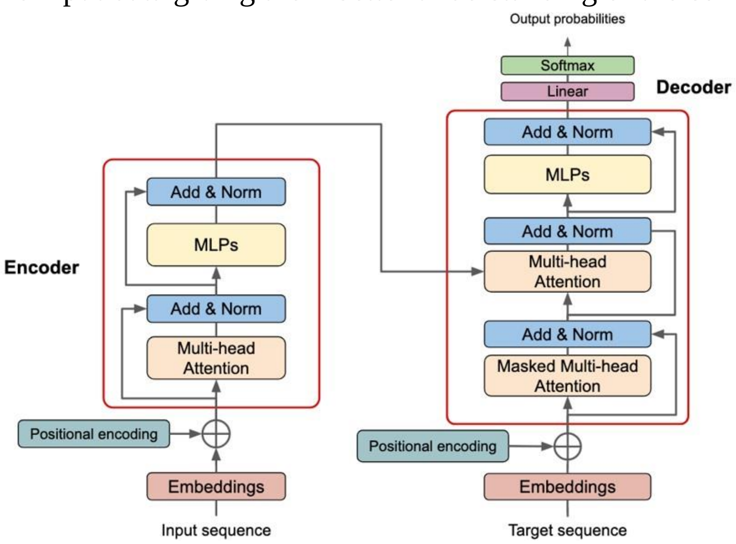

Figure 2: Diagram of Transformer architecture

Figure 2 illustrates the transformer architecture. The input sequence is converted into vector embeddings and positional embeddings to understand the order of the input words. The encoder of the transformer consists of two modules, one is multi-head attention, and the other is feed forward neural network. Multi head attention is where the model learns to pay attention to some words which are related to each other to understand the context. Feedforward is the fully connected neural network layer. After each layer, normalization step is performed, and output is connected back to its input. The output of last encoder with a rich contextual meaning goes to decoder component.

4.2 mBERT and DistilmBERT Architecture

The mBERT and DistilmBERT models contain only an encoder component of the original transformer architecture. The BERT and mBERT models are similar in architecture but the only difference is training data. The mBERT model is built on 12 transformer-encoder blocks, whereas DistilmBERT is built on only 6 transformer-encoder blocks, thus reducing its size to half of mBERT.

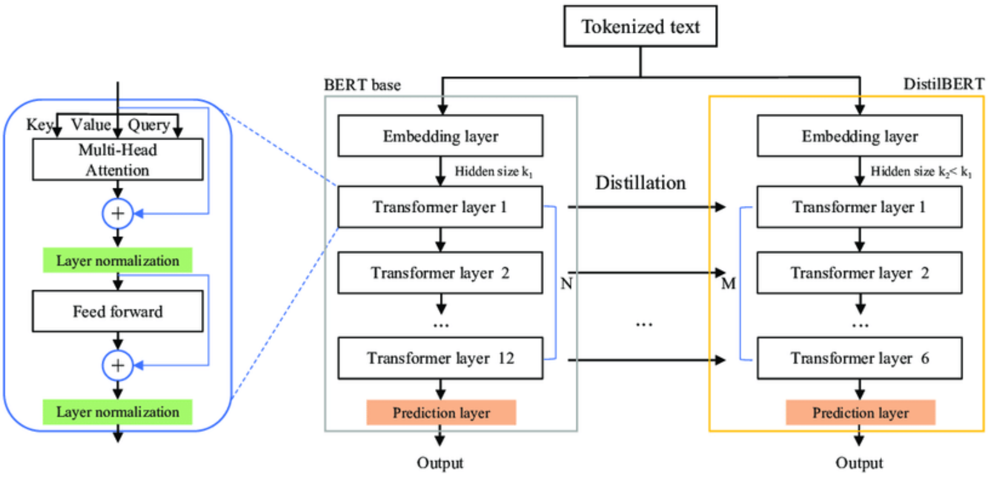

Figure 3: Architecture of BERT and DistilBERT

The architecture of BERT and DistilBERT is illustrated in Figure 3. The input text in our case Urdu is undergone embedding process where three operations are performed. First are token embeddings which convert Urdu words into vector representations. Second is segment embeddings which helps distinguish between two sentences to understand context of words in a sentence. Third is position embeddings which encode the position of each word in sentence. Input embeddings are broken into Query, Key and Value vectors. These vectors are computed with weights which models have learnt during training. Input embeddings are fed to attention layers where models learn which words are important and relevant to distinguish between fake and real news.

4.3 mT5 Architecture

Unlike BERT, T5 follows encoder-decoder configuration. The architecture of T5 and mT5 is similar but the only difference is training data and language scope. The developers evaluated the model on different benchmark tasks and found strong performances (Raffel et al., 2020). Figure 4 illustrates the unified framework of T5 model showing translation (denoted by green), linguistic acceptability (denoted by red), sentence similarity (denoted by yellow), and document summarization (denoted by blue). In T5 architecture, normalization is applied before each attention and neural network layer instead of after, unlike in original transformer architecture. T5 uses relative positional embeddings as opposed to absolute positional embeddings in transformer architecture, providing better generalization over various tasks.

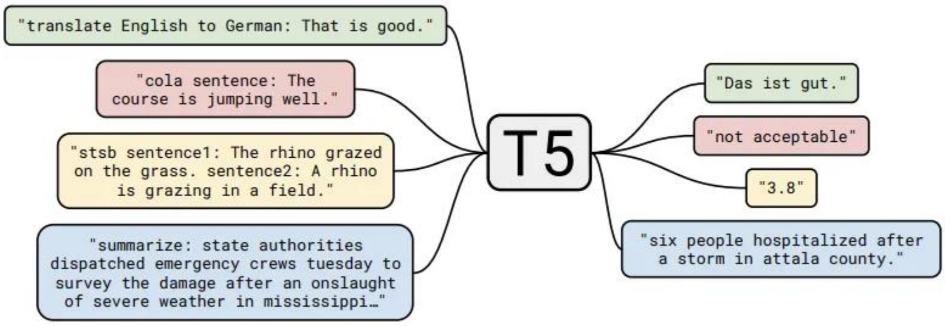

Figure 4: Diagram of T5 framework

5 Implementation

This section provides details on the different steps performed in this project. Starting from dataset selection to evaluation process, as depicted by Figure 5.

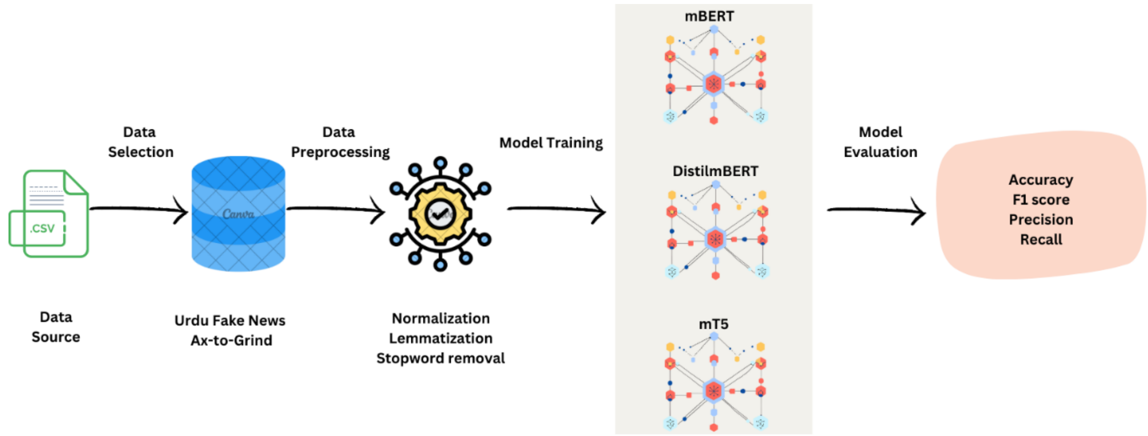

Figure 5: Implementation workflow for Urdu fake news detection system

5.1 Environmental Setup

This research utilizes Google Collaboratory which is cloud-based Jupyter notebook environment. It provides free access to powerful computational resources such as Tesla T4 GPU which is used in this project to build the fake news detection model. Tesla T4 GPU is based on Turing architecture and is much faster than CPU. For coding, Python programming language is used as it provides powerful libraries for language models and preprocessing textual data. Various python libraries are used to perform different steps in the coding section. NumPy and Pandas libraires are used for numerical computation and data manipulation. Matplot.lib is used for plotting graphs. Scikit learn library is used for train test split and evaluation metrics. Hugging Face[3] library is used for transformer model's implementation and PyTorch[4] is used as deep learning framework. More details on different python libraries, their versions and system configuration are provided in the configuration manual.

[3] https://huggingface.co/docs/hub/en/transformers
[4] https://pytorch.org/hub/huggingface_pytorch-transformers/

5.2 Urdu Fake News Dataset

As mentioned earlier, the dataset used in this research work is downloaded from GitHub repository[5] provided by Harris et al. (2023). This dataset contains a total 10,083 Urdu news items, comprising of 5053 fake news and 5030 from fifteen domains from 2017-2023. The dataset is balanced in terms of two classes, fake and real news as can be seen in the bar plot of Figure 6 where fake news is denoted by blue and true news is denoted by red.

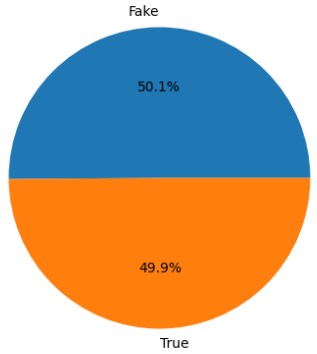

Figure 6: Distribution of news by class

S#	News Item	Label
1	ٹی ٹی پی نے پنجاب حکومت کے ہیلی کاپٹر کے عملے کو یرغمال بنانے کی تردید کی ہے۔ (TTP has denied taking the crew of the Punjab government helicopter hostage.)	Fake
2	مارک زکربرگ سیاست میں آنے کا سوچ رہے ہیں۔ (Mark Zuckerberg is thinking of getting into politics.)	Fake
3	فریدہ جلال نے اپنی موت کی افواہوں پر تنقید کی۔ (Farida Jalal criticized the rumors of her death.)	Fake
4	جعلی خبریں :پاپ اسٹار حدیقہ کیانی نے جعلی منشیات کے بارے میں برطانوی ویب سائٹ پر تنقید کی۔ (Fake news: Pop star Hadiqa Kayani slams British website about fake drugs)	Fake
5	صنم ماروی نے میڈیا پر گردش کرنے والی زیادتی اور ڈکیتی کی کوشش کی افواہوں کی تردید کی۔ (Sanam Marvi denied the rape and attempted robbery rumors circulating in the media.)	Fake

Table 2: Fake news content from dataset with English translation

S#	News Item	Label
1	AAP دہلی میں جلد ہی 24×7 مال، تجارتی علاقوں میں کھانے پینے کی جگہیں ہو سکتی ہیں اگر کجریوال کی واپس آتی ہے (Delhi may soon have 24x7 malls, eateries in commercial areas if Kejriwal's AAP returns)	Real
2	بی جے پی مسلسل ہاردک پٹیل کو ہراساں کر رہی ہے :کانگریس لیڈر پرینکا گاندھی واڈرا (BJP is constantly harassing Hardik Patel: Congress leader Priyanka Gandhi Vadra)	Real
3	کیرالہ اور پنجاب کے بعد مہاراشٹرا بھی سی اے اے کے خلاف قرارداد پر غور کر رہا ہے۔ (After Kerala and Punjab, Maharashtra is also considering a resolution against CAA.)	Real
4	یوپی کے سی ایم یوگی آدتیہ ناتھ کے قریبی ساتھی سنیل سنگھ سماج وادی پارٹی میں شامل ہو گئے۔ (UP CM Yogi Adityanath's close aide Sunil Singh has joined the Samajwadi Party.)	Real
5	مودی حکومت کے وزراء کشمیری سیاست دانوں سے مل سکتے ہیں جو ’ابھرتے ہوئے محاذ ‘کا حصہ ہیں۔ (Modi government ministers may meet Kashmiri politicians who are part of 'emerging front'.)	Real

Table 3: Real news content from dataset with English translation

[5] https://github.com/Sheetal83/Ax-to-Grind-Urdu-Dataset

5.3 Data Preprocessing

The raw data contains irrelevant information which do not add to the learning of model. Below are some steps performed on raw dataset which help model to efficiently handle text data.

5.3.1 Normalization

Normalization is performed to standardize the Urdu text by removing diacritics and normalizing character variation. One example is Arabic character 'آ' and other Urdu character 'آ' is normalized to one character considering the context they are used in. urdu_text_processing.normalize() command from LughaatNLP library is used to perform this normalization technique on our Urdu dataset.

5.3.2 Lemmatization

Lemmatization technique is performed to get the root word of Urdu text data as it helps lessen the input space. The lemmatization is more accurate than stemming in terms of Urdu text as stemming only removes the prefixes and suffixes from a word however Urdu text also contain infixes which are present between root word (Hafeez et al., 2023). Considering the linguistic intricacies, this step is performed. The method urdu_text_processing.lemmatize_sentence() from LughaatNLP Library is used to lemmatize Urdu text. Figure 7 shows the lemmatized output of Urdu sentence which converts the words into their base form. The Urdu text مجھے اردو کے کتابچہ پڑھنا اچھا لگتا ہے (I like reading Urdu books) is transformed into separate words with their root form such as پڑھنا (reading) to پڑھ (read).

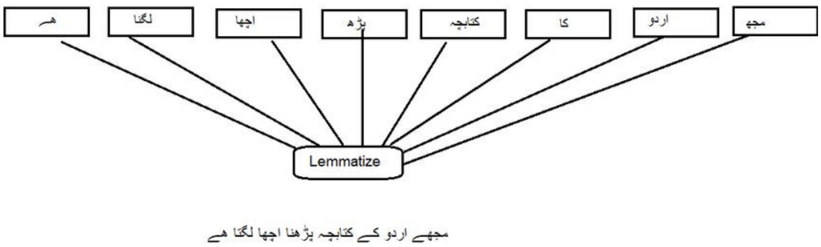

Figure 7: Example of lemmatization of Urdu text

5.3.3 Stop words removal

Stop words in Urdu language are conjunctions or phrases that grammatically limit a sentence and have very low semantic and contextual importance (Harris et al., 2023). Therefore, these stop words are eliminated to reduce the noise from the dataset using LughaatNLP library's function urdu_text_processing.remove_stopwords(). Figure 8 shows the word cloud of stop words of whole dataset. The bolder font size, the more frequent it occurs in data such as کے (that), اور (and) and ہے (is) are bigger in font size and repeat multiple times in our text data. Python's WordCloud library is used to generate word cloud. The list of Urdu stop words is downloaded from GitHub[6] which is compared with tokenized words of Urdu to remove stop words. Tokenization was performed with Spacy library as it provides support for Urdu language. Figure 9 shows the word cloud after removing stop words. Most frequent words which appear in news data are now پاکسان (Pakistan), ایک (one) and ٹرمپ (Trump).

Figure 8: Word cloud before removing stop words

Figure 9: Word cloud after removing stop words

5.3.4 Special character removal

The Urdu text contains some special character and non-word character such as #, $, %, & and punctuations symbols. These special words do not provide any semantic meaning and thus are removed from data. Regular expression (re) is used to remove all these characters and keep only Urdu letters (ا-ی) and numbers (۰-۹) for Urdu specific filtering.

5.3.5 Textual Data Analysis

Table 4 lists the various analysis performed on Urdu dataset. Fake news items contain longer sequence as compared to real news with average character count of 528 and word count of 115 whereas for real news the average character length is 171 and word count 34 per news item. The graphical illustration of word count by class label is illustrated in Figure 10 where red denotes true news and blue denotes fake news. It can be observed that distribution is heavily skewed towards right indicating that majority of text has shorter word counts. It is important as it helps us decide the value of maximum length of tokens which will be fed to the model.

[6] https://github.com/SyedMuhammadMuhsinKarim/Urdu-Stop-Words

Feature	Fake	Real
Average characters count per news	528.6	171.7
Minimum text length	13	25
Maximum text length	5224	835
Average word count per news	115.8	34
Minimum word count	3	4
Maximum word count	1188	190
Total Entries	5053	5030

Table 4: Urdu text analysis

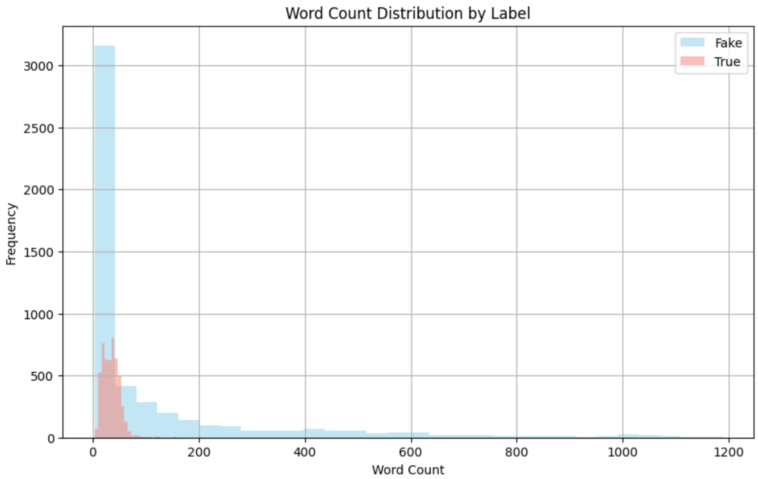

Figure 10: Word count plot by class label

5.4 Model Implementation

This section provides details on the implementation of the pre-trained transformer models. The cleaned text data is split into training, validation and testing set with a ratio of 60:20:20 as suggested by Harris et al. (2023) using Python's scikit-learn library. Table 5 lists the proportion of news items for both fake and real news after splitting data into train, validation and test set. It can be observed that training set contains a total of 6,049 news items, out of which fake news are 3,031 and real news are 3,018. It means that model will be trained on balanced dataset and will not have bias to one class. The test set is used to evaluate model performance on unseen data.

Data	Fake News	Real News	Total
Train set	3,031	3,018	6,049
Validation set	1,011	1,006	2,017
Test set	1,011	1,006	2,017
Total	5,053	5,030	10,083

Table 5: The class-wise number of samples after data split

5.4.1 mBERT Implementation

Hugging face library provides easy implementation of various transformer models by importing transformer library into the notebook file. Each model has its own tokenizer which is required to tokenize the textual data. In this implementation, BertTokenizer is used to tokenize the input text and BertForSequenceClassification is used to initialize the pretrained "bert-base-multilingual-cased" model. The maximum length of token vector is set to 128 because the average word count from Urdu text analysis is 115 for fake news and 34 for real news, as listed in Table 4. DataLoader package is used for efficient data batching and shuffling of training and validation data into models. The model is trained for ten epochs with training data and evaluated on validation data. The epoch number was set to 10, the batch size is set to 16, the learning rate is set to 2xe-5 and AdamW optimizer is used to optimize the loss function. These are hyperparameters and are set as suggested in previous studies (Harris et al., 2023). AdamW optimizer as compared to Adam optimizer provides decoupling weight decay from the gradient update which allows for better generalization of models such as BERT (Loshchilov, 2017). Many studies tune the hyperparameters to optimize model performance such as varying learning rate, increasing epoch number and trying with different batch size. However, in this study these hyperparameters are kept fixed due to constraints of computational resources. For every epoch training loss and validation accuracy is calculated and the training progress is visualized using the tqdm library available in python.

5.4.2 DistilmBERT Implementation

The implementation of DistilmBERT is same as that of mBERT base model. Imported from hugging face, DistilBertTokenizer tokenizes the input text for model. The DistilmBERT model “distilbert-base-multilingual-cased” from hugging face is initialized with DistilBertForSequenceClassification. Numer of labels is set to 2 when initializing the model as model predicts class label either fake or true news. All the hyperparameters are kept the same as that of mBERT implementation to compare the performance between them.

5.4.3 mT5 Implementation

The small version of mT5 model is used for implementation as it is suitable for less computational demand. From hugging face transformer library, mT5 small model and its tokenizer are imported. MT5Tokenizer is used to tokenize the Urdu input text sequence into vector embeddings. The pretrained “"google/mt5-small" model is initialized with MT5ForConditionalGeneration which generates output in text form as per condition. In our case, the output is a label with fake and true class which will be generated as text. The function model.generate() is used to generate the output sequence with a max_length of 10 tokens because the output labels are true and fake strings which are only four characters in length. Dataloaders are created for train and validation data using DataLoader() with batch size of 8. Initially batch size was kept 16 same as that of other models in this research but is showed error “OutOfMemoryError” meaning that allocated RAM memory was consumed. After reading about this error, the recommendation was to decrease batch size[7] and therefore batch size of 8 was used and memory error was resolved (Sabry et al., 2022). Learning rate and epoch were set to10 same as other models but at 10th epoch, training and validation accuracy were 49% and 51% respectively. Therefore, number of epochs was increased to 20 to check model performance.

6 Evaluation and Results

This section provides analysis of the results obtained from experiments conducted in this study. The results are presented using evaluation metrics to assess the performance of each model.

6.1 Experiment 1: mBERT model for Urdu fake new detection

The Figure 11 shows the plot of training loss (red line) and validation accuracy (blue line) over ten epochs for mBERT implementation. After each epoch, training loss is decreasing showing that model is learning from labelled data. Figure 12 shows the classification report which provides insights into evaluation metrics for each class. These evaluation metrics are tabulated in Table 6. Figure 13 depicts the confusion matrix which shows the number correct and incorrect predictions by each class. The mBERT has achieved an accuracy of 89%, indicating that 89% of times model's predictions were correct. Precision value of 91% for fake class shows that out of total 960 predictions by model as fake news, 876 news were indeed fake news. Whereas precision value of 87% for real news shows that out of total 1057 predictions by model as real news, 822 news were indeed real news. Recall value of 87% for fake news shows that out of all 1,011 fake news in test data, model was able to correctly predict 876 news as fake news. Recall value of 92% for real news shows that out of all 1006 real news in test data, model was able to correctly predict 922 news as real news. The mBERT model achieves better precision for fake news and better recall for real news. These insights are useful they show ability of model to avoid misclassifying real news as fake (better precision) and ensures more real news are detected (better recall).

Figure 11: Training loss and validation accuracy for mBERT

[7] https://www.restack.io/p/gpu-computing-answer-google-colab-t4-gpu-limit-cat-ai

```
Classification Report:
              precision    recall  f1-score   support

           0       0.91      0.87      0.89      1011
           1       0.87      0.92      0.89      1006

    accuracy                           0.89      2017
   macro avg       0.89      0.89      0.89      2017
weighted avg       0.89      0.89      0.89      2017
```

Figure 12: Classification report of mBERT implementation

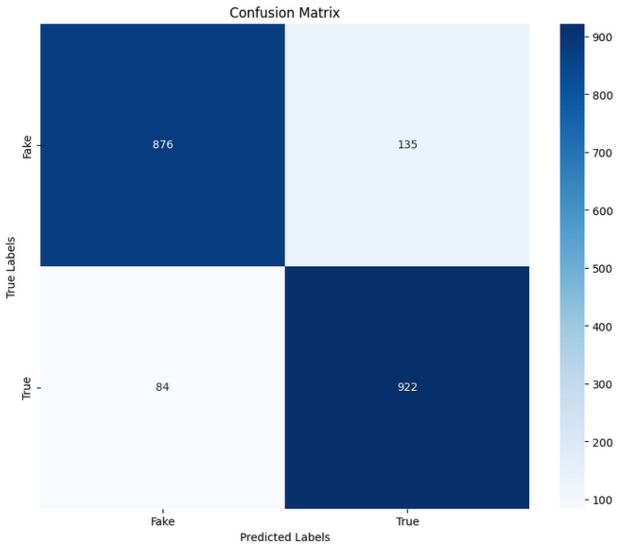

Figure 13: Confusion matrix for mBERT implementation

Metric	Fake News	Real News
Accuracy (%)	89	89
Precision (%)	91	87
Recall (%)	87	92
F1-Score	89	89

Table 6: Evaluation metrics for mBERT implementation

6.2 Experiment 2: DistilmBERT for Urdu fake new detection

The training loss and validation accuracy of DistilmBERT model over the ten epochs is depicted by Figure 14. The diagram suggests that model learns to differentiate between fake and real news from training data as training loss decreases over the epochs. Performance of model on validation set indicates no sign of overfitting as trend shows the increasing validation accuracy up to 90% over ten epochs. However, there is slight decrease in validation accuracy after epoch 6 but is minimal. Figure 15 shows the classification report which provides metric values for both classes fake (Class 0) and real (Class 1) news. The evaluation metrics are also listed in tabular format in Table 7. Figure 16 shows the confusion matrix indicating number of samples misclassified or correctly classified by the DistilmBERT model. Out of total 2,017 test data, DistilmBERT model predicted 1792 cases correctly on both classes and 225 news were misclassified as fake or real, giving us the accuracy of 89%. Precision value of 88% for fake class shows that out of total 1,040 predictions by model as fake news, 913 news were indeed fake news. Whereas precision value of 90% for real news shows that out of total 977 predictions by model as real news, 879 news were indeed real news. Recall value of 90% for fake news shows that out of all 1,011 fake news samples, model was able to correctly predict 913 news as fake news. Recall value of 87% for real news shows that out of all 1006 real news samples, model was able to correctly predict 879 news as real news. F1-score (89%) is a balanced score which is same for both the classes. The results of DistilmBERT indicate overall strong performance on all metrics. The model has shown better precision for real news and better recall for fake news. That means model is less likely to overlook fake news and misclassify it as real news (better precision) and identifies fake news more effectively than real news (better recall) which is critical in combating misinformation

Figure 14: Training loss and validation accuracy for DistilmBERT

Classification Report:

	precision	recall	f1-score	support
0	0.88	0.90	0.89	1011
1	0.90	0.87	0.89	1006
accuracy			0.89	2017
macro avg	0.89	0.89	0.89	2017
weighted avg	0.89	0.89	0.89	2017

Figure 15: Classification report for DistilmBERT

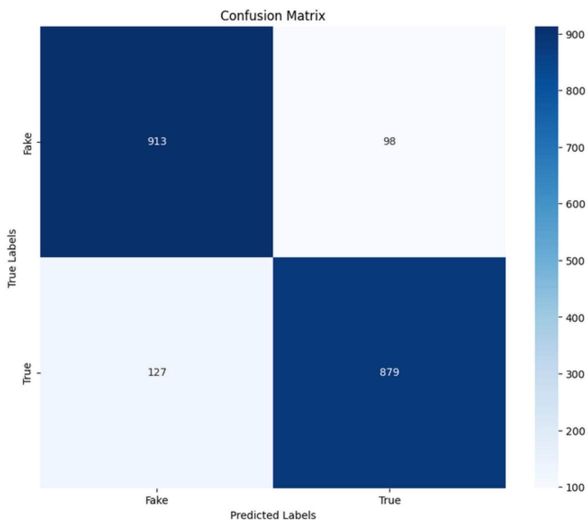

Figure 16: Confusion matrix for DistilmBERT

Metric	Fake News	Real News
Accuracy (%)	89	89
Precision (%)	88	90
Recall (%)	90	87
F1-Score	89	89

Table 7: Evaluation Metrics for DistilmBERT

6.3 Experiment 3: mT5 for Urdu fake new detection

In Figure 17, training loss (blue) and validation accuracy (green) of mT5 model is visualized for 20 epochs. The spikes and dips in validation accuracy suggest potential overfitting which is adjusted as training progresses. The confusion matrix is shown in Figure 18. The classification report is shown in Figure 19. Table 8 lists evaluation metrics for each class.

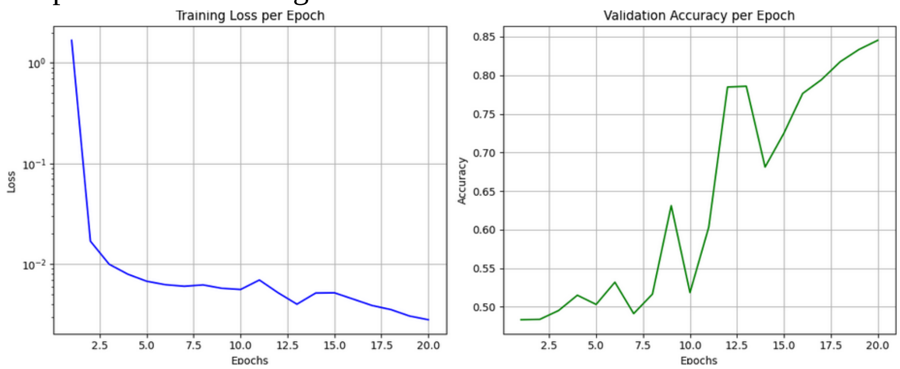

Figure 17: Training loss and validation accuracy for mT5

The mT5 model has achieved an accuracy of 84%, F1-score of 85% for fake news and 83% for real news. Precision value (81%) for fake class shows that out of total 1,153 predictions by model as fake news, 935 news were correct. Whereas precision value of 88% for real news shows that out of total 864 predictions by model as real news, 763 news were indeed real news. Recall value of 90% for fake news shows that out of all 1,036 fake news samples, model was able to correctly predict 935 news as fake news. Recall value of 78% for real news shows that out of all 981 real news samples, model was able to correctly predict 763 news as real news.

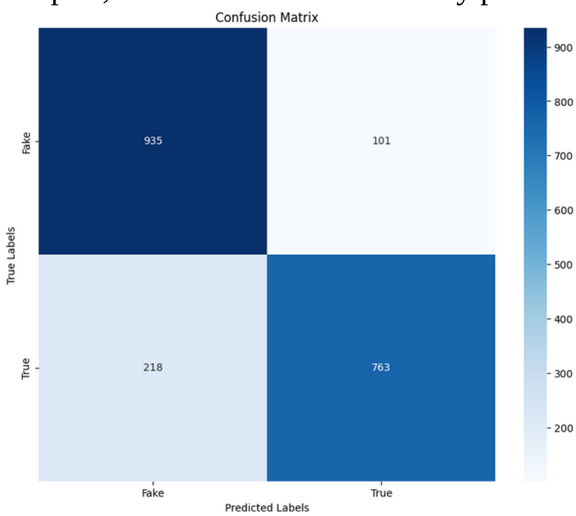

Figure 18: Confusion matrix of mT5 implementation

```
Classification Report:
              precision    recall  f1-score   support

        Fake       0.81      0.90      0.85      1036
        True       0.88      0.78      0.83       981

    accuracy                           0.84      2017
   macro avg       0.85      0.84      0.84      2017
weighted avg       0.85      0.84      0.84      2017
```

Figure 19: Classification report of mT5 implementation

Metric	Fake News	Real News
Accuracy (%)	84	84
Precision (%)	81	88
Recall (%)	90	78
F1-Score	85	83

Table 8: Evaluation metrics of mT5

6.4 Discussion

The models mBERT and DistilmBERT have demonstrated similar performance in terms of accuracy and F1-score. This is because DistilBERT retains 97% language understanding capabilities as compared to BERT model despite being smaller in size. However, the performance of both these models in terms of precision and recall is opposite. Table 9 lists the comparative analysis of evaluation metrics for all experiments for both classes. The mBERT has the highest precision for fake news, whereas DistilmBERT has the highest precision for real news. In other words, mBERT is more cautious to label news as fake news to avoid mistakenly classifying real news as fake news, thereby minimizing the risk of spreading misinformation. On the other hand, DistilmBERT is the most effective in predicting fake news accurately as compared to real news. The mT5 model, as compared to other models, lags in overall performance with lowest accuracy, precision and F1-score. However, it has good recall for fake news but very low recall for real news, indicating that mTt5 model is more inclined to classifying news as fake therefore increasing the chance of spreading misinformation if real news is classified as fake.

Evaluation Metrics	**mBERT**		**DistilmBERT**		**mT5**	
	Fake	**Real**	**Fake**	**Real**	**Fake**	**Real**
Accuracy (%)	89	89	89	89	84	84
Precision (%)	**91**	87	88	**90**	81	88
Recall (%)	87	**92**	**90**	87	90	78
F1-Score	89	89	89	89	85	83

Table 9: Comparison of evaluation metrics for models

The findings of this research suggest that DistilmBERT offers promising results for fake news detection in Urdu, enabling its use in resource-efficient environments, The other finding is that mT5 model has underperformed for fake news detection. This might be because this model is pre-trained on text generation tasks and is suited for NLP tasks where text is generated as output such as summarization and question-answering. Moreover, mT5 requires more computational resources due to its encoder-decoder architecture. The accuracy of the model is less as compared to previous study by Harris et al. (2023) as can be seen in Table 1 of comparative literature analysis in Section 2. In pervious study, authors performed ensemble technique on three transformer models. In comparison to that, this study demonstrates promising results from just standalone DistilmBERT transformer model requiring significantly less computational resources as compared to running three large language models simultaneously.

7 Conclusion and Future Work

The aim of this study is to investigate the performance of transformer models for fake news classification in low resource language Urdu. Three multilingual variants of transformer models mBERT, DistilmBERT and mT5 are employed in this research. Implementation results show that both mBERT and DistilmBERT models outperfdorm mT5, suggesting that BERT based encoder-only transformer models are more robust for classification of fake news. The DistilmBERT and mT5 models are first explored in this study for fake news detection in Urdu. This study also finds that both mBERT and DistilmBERT showed comparable performance overall. DistilmBERT, despite lightweight and efficient model as compared to mBERT still delivers solid results, making it an appropriate choice for resource-constrained environments such as deploying models in real world applications. This study answers the research question by demonstrating promising results for fake news classification in low resource language Urdu using advanced transformer-based architecture. This work sets a foundational framework to advance the research in less-explored domain of fake news detection in low resource languages.

Future work can include using ensemble techniques such as voting mechanism on DistilmBERT and other transformer models to improve classification accuracy by leveraging the strength of individual models as suggested in literature. For instance, an ensemble model developed with DistilmBERT and mBERT can capitalize on the efficiency of DistilmBERT and the robust contextual understanding of mBERT. Also, different hyperparameters such as batch size, number of epochs, split ratio and learning rates can be tried to optimize model performance. Due to time and compute constraints, this study has kept fixed hyperparameters to simplify experimentations of different models. The usage of fixed hyperparameters might have limited the model's ability to uncover subtle patterns in Urdu dataset such as colloquialisms or informal language variations. Therefore, future study can consider fine-tuning these parameters to enhance model's ability to adopt to the nuances of Urdu dataset, leading to a more generalized and efficient fake news detection model.

This research can potentially help media organizations, government institutions, journalists and other stakeholders to address the challenge of countering fake news in growing population of social media users. The implications of this work can be significant on society and individuals. By improving the detection of fake news in low resource language, this research can contribute to develop more social trust and help reduce the societal divisions caused by misinformation.

References

Aïmeur, E., Amri, S. and Brassard, G. 2023. Fake news, disinformation and misinformation in social media: a review. *Social Network Analysis and Mining*. **13**(1), p.30.

Akhter, M.P., Zheng, J., Afzal, F., Lin, H., Riaz, S. and Mehmood, A. 2021. Supervised ensemble learning methods towards automatically filtering Urdu fake news within social media. *PeerJ Computer Science*. **7**, p.e425.

Alghamdi, J., Luo, S. and Lin, Y. 2023. A comprehensive survey on machine learning approaches for fake news detection. *Multimedia Tools and Applications*. **83**, pp.1–59.

Amjad, M., Sidorov, G. and Zhila, A. 2020. Data augmentation using machine translation for fake news detection in the Urdu language *In*: *LREC 2020 - 12th International Conference on Language Resources and Evaluation, Conference Proceedings*.

Amjad, M., Sidorov, G., Zhila, A., Gomez Adorno, H., Voronkov, I. and Gelbukh, A. 2020. “Bend the truth”: Benchmark dataset for fake news detection in Urdu language and its evaluation. *Journal of Intelligent & Fuzzy Systems*. **39**, pp.1–13.

Arshad, U., Malik, K.I. and Arooj, H. 2022. Urdu News Content Classification Using Machine Learning Algorithms. *Lahore Garrison University Research Journal of Computer Science and Information Technology*.

Azizah, S.F.N., Cahyono, H.D., Sihwi, S.W. and Widiarto, W. 2023. Performance Analysis of Transformer Based Models (BERT, ALBERT, and RoBERTa) in Fake News Detection *In*: *2023 6th International Conference on Information and Communications Technology (ICOIACT).*, pp.425–430.

Balaji, N.N.A. and Bharathi, B. 2020. SSNCSE_NLP@Fake news detection in the Urdu language (UrduFake) 2020 *In*: *CEUR Workshop Proceedings*.

Baruah, A., Das, K., Barbhuiya, F.A. and Dey, K. 2020. Automatic Detection of Fake News Spreaders Using BERT *In*: *Conference and Labs of the Evaluation Forum* [Online]. Available from: https://api.semanticscholar.org/CorpusID:225072699.

Bhawal, S. and Roy, P.K. 2021. Fake News Detection in Urdu Language using BERT *In*: *Fire* [Online]. Available from: https://api.semanticscholar.org/CorpusID:251322545.

De, A., Bandyopadhyay, D., Gain, B. and Ekbal, A. 2022. A Transformer-Based Approach to Multilingual Fake News Detection in Low-Resource Languages. *ACM Transactions on Asian and Low-Resource Language Information Processing*. **21**, pp.1–20.

Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding *In*: *North American Chapter of the Association for Computational Linguistics* [Online]. Available from: https://api.semanticscholar.org/CorpusID:52967399.

Farooq, M.S., Naseem, A., Rustam, F. and Ashraf, I. 2023. Fake news detection in Urdu language using machine learning. *PeerJ Computer Science*. **9**, p.e1353.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B. and Lazer, D. 2019. Fake news on Twitter during the 2016 U.S. presidential election. *Science*. **363**(6425), pp.374–378.

Hafeez, R., Anwar, M.W., Jamal, M.H., Fatima, T., Espinosa, J.C.M., López, L.A.D., Thompson, E.B. and Ashraf, I. 2023. Contextual Urdu Lemmatization Using Recurrent Neural Network Models. *Mathematics*. **11**(2).

Harris, S., Liu, J., Hadi, H.J. and Cao, Y. 2023. Ax-to-Grind Urdu: Benchmark Dataset for Urdu Fake News Detection *In*: *2023 IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*., pp.2440–2447.

Jwa, H., Oh, D., Park, K., Kang, J.M. and Lim, H. 2019. exBAKE: Automatic Fake News Detection Model Based on Bidirectional Encoder Representations from Transformers (BERT). *Applied Sciences*. **9**(19).

Kareem, I. and Awan, S.M. 2019. Pakistani Media Fake News Classification using Machine Learning Classifiers *In*: *2019 International Conference on Innovative Computing (ICIC)*., pp.1–6.

Kishwar, A. and Zafar, A. 2022. Fake news detection on Pakistani news using machine learning and deep learning. *Expert Systems with Applications*. **211**, p.118558.

Lin, N., Fu, S. and Jiang, S. 2020. Fake News Detection in the Urdu Language using CharCNN-RoBERTa *In*: *Fire* [Online]. Available from: https://api.semanticscholar.org/CorpusID:232314734.

Loshchilov, I. 2017. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*.

Panda, S. and Levitan, S.I. 2021. Detecting Multilingual COVID-19 Misinformation on Social Media via Contextualized Embeddings *In*: A. Feldman, G. Da San Martino, C. Leberknight and P. Nakov, eds. *Proceedings of the Fourth Workshop on NLP for Internet Freedom: Censorship, Disinformation, and Propaganda* [Online]. Online: Association for Computational Linguistics, pp.125–129. Available from: https://aclanthology.org/2021.nlp4if-1.19.

Qazi, M., Khan, M.U.S. and Ali, M. 2020. Detection of Fake News Using Transformer Model *In*:, pp.1–6.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W. and Liu, P.J. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of machine learning research*. **21**(140), pp.1–67.

Rafique, A., Rustam, F., Narra, M., Mehmood, A., Lee, E. and Ashraf, I. 2022. Comparative analysis of machine learning methods to detect fake news in an Urdu language corpus. *PeerJ Computer Science*. **8**, p.e1004.

Rahman, S., Sharif, O., Das, A., Afroze, S. and Hoque, M. 2022. FaND-X: Fake News Detection using Transformer-based Multilingual Masked Language Model *In*:, pp.153–158.

Ranjan, V. and Agrawal, P. 2022. Fake News Detection: GA-Transformer And IG-Transformer Based Approach. *2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*., pp.487–493.

Reddy, S.M., Suman, C., Saha, S. and Bhattacharyya, P. 2020. A GRU-based Fake News Prediction System: Working Notes for UrduFake-FIRE 2020 *In*: *Fire* [Online]. Available from: https://api.semanticscholar.org/CorpusID:232313568.

Sabry, S.S., Adewumi, T., Abid, N., Kovács, G., Liwicki, F. and Liwicki, M. 2022. Hat5: Hate language identification using text-to-text transfer transformer *In*: *2022 International Joint Conference on Neural Networks (IJCNN)*. IEEE, pp.1–7.

Sanh, V., Debut, L., Chaumond, J. and Wolf, T. 2019. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *ArXiv*. **abs/1910.01108**.

Shariff, M., Thoms, B., Isaacs, J.T. and Vakilian, V. 2022. Approaches in Fake News Detection: An Evaluation of Natural Language Processing and Machine Learning Techniques on the Reddit Social Network. *Artificial Intelligence and Applications*.

Shu, K., Sliva, A., Wang, S., Tang, J. and Liu, H. 2017. Fake News Detection on Social Media: A Data Mining Perspective. *ACM SIGKDD Explorations Newsletter*. **19**.

Sun, W., Cai, Z., Li, Y., Liu, F., Fang, S. and Wang, G. 2018. Data Processing and Text Mining Technologies on Electronic Medical Records: A Review. *Journal of Healthcare Engineering*. **2018**, pp.1–9.

Tahir, B. and Mehmood, A. 2022. UBERT22: Unsupervised Pre-training of BERT for Low Resource Urdu Language *In*:, pp.1–6.

Ullah, F., Zamir, M., Arif, M., Ahmad, M., Felipe-Riveron, E. and Gelbukh, A. 2024. Fida @DravidianLangTech 2024: A Novel Approach to Hate Speech Detection Using Distilbert-base-multilingual-cased. , pp.85–90.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. ukasz and Polosukhin, I. 2017. Attention is All you Need *In*: I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett, eds. *Advances in Neural Information Processing Systems* [Online]. Curran Associates, Inc. Available from: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

Vosoughi, S., Roy, D. and Aral, S. 2018. The spread of true and false news online. *Science*. **359**(6380), pp.1146–1151.

Xue, L. 2020. mt5: A massively multilingual pre-trained text-to-text transformer. *arXiv preprint arXiv:2010.11934*.