

# DeFactify

PROJECT PROPOSAL

MARCH 20, 2025



Ivet Kalcheva

Version 0.1

(Draft)

# Contents

---

Introduction.....	3
What? .....	3
Why? .....	3
Who? .....	3
When? .....	3
How? .....	4
Domain Understanding .....	4
Research Question .....	4
Exploratory Research.....	4
Research Methods .....	4
Analytic Approach .....	5
Target Variable .....	5
Type of Problem.....	5
Potential Machine Learning Models .....	5
Defining success .....	5
Data Requirements.....	6
Define Objectives.....	6
Data Characteristics .....	6
Data Sources .....	6
Data Legality and Ethics .....	6
Data Diversity .....	6
Version Control .....	6
Iterative Process .....	6
Conclusion .....	7

# Introduction

---

## What?

The goal of this project is to develop a machine learning model that detects fake news by analysing the sources and recognising linguistic patterns in news article through a link. The artificial intelligence must accurately identify real articles and differentiate them from fake ones, while also providing insights into why an article may be flagged as false. This will help users in identifying unreliable sources of information while surfing online.

## Why?

Fake news has become an increasing concern, especially in recent years due to the growing influence of artificial intelligence and social media. In the past, there were many regulations governing journalism. However, in today's digital age, the internet allows anyone to publish, share and consume information and news with little to no oversight or professional standards. This trend has led to much misinformation, social disruption and, at times, the use of fake news as a tool for spreading propaganda. This issue is particularly alarming for many older adults, who often trust blindly the information they read on social media. Therefore, a tool that warns them to the dangers of fake news would be of great importance.

## Who?

- ✦ **Stakeholder:** Yulian Kalchev – a father and victim of fake news on social media.
- ✦ **General Public** would benefit from this assistance, during their online presence.

## When?

End Date	Week 5	Week 7	Week 8	Week 9	Week 10
Iteration 0					
Iteration 1					
Iteration 2					
Iteration 3					
Iteration 4					

- ✦ **Iteration 0:** Choose an appropriate idea, draft the first version of the project proposal, data collection of example news, environment setup and implementation.
- ✦ **Iteration 1:** Implement feedback from iteration 0 and improve the implementation.
- ✦ **Iteration 2:** Implement feedback from iteration 1 and improve the implementation.
- ✦ **Iteration 3:** Implement feedback from iteration 2 and complete optimisation.
- ✦ **Iteration 4:** Last iteration with final touch-ups, putting a focus on the presentation.

## How?

The final product will be a web application, where users can paste a news article link to receive a classification ("*Truthful*" or "*Fake*") alongside a summary of key indicators that influenced the decision. The model will be updated periodically to incorporate new trends in fake news, so it stays accurate.

## Domain Understanding

---

### Research Question

What linguistic and structural features distinguish fake news from real news?

### Exploratory Research

To ensure high prediction accuracy, the research will focus on:

- ✦ Identifying fully truthful sources of news.
- ✦ Recognising linguistic differences of fake and real news (e.g., sensational language).
- ✦ Analysing headline patterns and their influence on perceived authenticity (e.g., clickbait<sup>1</sup>).
- ✦ Investigating metadata<sup>2</sup> (e.g., publication date, author reputation).

### Research Methods

- ✦ **Literature study:** Review prior studies on fake news detection.
- ✦ **Document Analysis:** Identify differences, analyse fake/real news articles and collect the data.
- ✦ **Model Evaluation:** Implement and evaluate models on accuracy and dataset.

---

**Clickbait<sup>1</sup>** - Clickbait stories use sensationalist headlines to grab attention and drive attention to the publisher's website usually at the expense of truth or accuracy.

**Metadata<sup>2</sup>** - Website metadata includes page titles, descriptions, keywords, etc. to help search engines rank web pages. Search engines interpret it to determine how relevant it is to a user's search query.

# Analytic Approach

---

## Target Variable

The target variable for this project could be represented through binary classification: either "*Fake*" (0) or "*Truthful*" (1). The model would be able to predict that based on the articles' source, linguistic features, headline patterns and metadata.

## Type of Problem

Classification problem, using supervised learning.

## Potential Machine Learning Models

- ✦ **K-Nearest Neighbours (KNN):** A simple and understandable model, suitable for initial testing with well-defined features.
- ✦ **Decision Trees:** Good for exploring feature importance and handling non-linear relationships, combining with feature engineering to test structural or linguistic cues.
- ✦ **Support Vector Machines (SVM):** Highly effective for text classification, especially with a well-defined margin between classes. Using kernel trick to deal with high-dimensional data.

## Defining success

Success will be defined by:

- ✦ **Prediction Accuracy:** High classification accuracy on the test dataset.
- ✦ **Robustness:** Generalisation across diverse datasets.
- ✦ **Stakeholder Satisfaction:** Positive feedback on usability and effectiveness.

# Data Requirements

---

## Define Objectives

To classify news articles as fake or real and provide explanations for classifications.

## Data Characteristics

- ✦ Text data from news articles (headlines, body text).
- ✦ Metadata such as publication date, author and source credibility.

## Data Sources

Scraping data on my own after extensive research.

## Data Legality and Ethics

- ✦ Ensure compliance with data privacy regulations (e.g., GDPR).
- ✦ Avoid misuse of data or spreading misinformation through the analysis.

## Data Diversity

Include articles from various domains and sources, including diverse topics (e.g., politics, health, entertainment, gossip) to improve generalisation.

## Version Control

Using GitHub for code and dataset versioning: **["DeFactify"]**

## Iterative Process

Continuous evaluations of model performance, refining preprocessing steps, updating dataset and incorporating constantly the given feedback.

# Conclusion

---

Overall, “*DeFactify*” targets the growing issue of misinformation by developing a machine learning algorithm for fake news detection. By analysing linguistic patterns and source credibility, the tool will help individuals identify unreliable information and promote critical thinking. The final product will be a user-friendly application, designed to fight against propaganda effectively.