ECE
ÉCOLE D'INGÉNIEURS
ENGINEERING SCHOOL

False notes on internet

-

Book of specification

Student :

Alex DONG – MSc 1 Data management and AI

Hamza LAMINE – MSc 1 Data management and AI

Narcis GEORGE – MSc 1 Data management and AI

RAJU Ajith – MSc 1 Sustainable and Energy

SAH Sanjeevkumar – MSc 1 Sustainable and Energy

Promo : 2023 – 2025

Tutors : M. Geoffroy LARDEUX – M. Lucas BORGNON – M. Daniel BURUIAN

Table of contents

Preamble

This document is a study paper leading to the MSc Diploma in Data Management and Artificial Intelligence. The use of its contents in no way engages the responsibility of the student or of the University Institute and its teachers.

Introduction

In today's increasingly digitalized world, online reviews play a crucial role in consumer decision-making. Whether buying products, booking services or choosing restaurants, reviews left by other users have a considerable influence on the choices made by Internet users. However, the proliferation of fake reviews poses a major threat to consumer confidence and the integrity of online platforms.

Fake reviews, often written by bots or sponsored by unscrupulous companies, can mislead users and distort perceptions of the real quality of products and services. They harm not only consumers, but also honest companies whose reputations are unfairly damaged.

Summary sheet

2023/2024 session	
Institution	**ECE Paris**

Project	False review on internet
Code	MSc #PPE23-P-238
Purpose of the project	The project will involve the design and implementation of a system based on artificial intelligence and natural language processing (NLP) to analyze online reviews and identify those that are likely to be false.
Project team :	Alex DONG – MSc 1 Data management and AI Hamza LAMINE – MSc 1 Data management and AI Narcis GEORGE – MSc 1 Data management and AI RAJU Ajith – MSc 1 Sustainable and Energy SAH Sanjeevkumar – MSc 1 Sustainable and Energy

Objective	The main aim of the "Detecting False Opinions on the Internet" project is to develop a technological solution capable of identifying and filtering out false opinions. This solution will aim to : • Strengthen consumer confidence by offering authentic and reliable reviews. • Protect honest businesses by ensuring that reviews of their products and services reflect real experiences. • Improve transparency on online review platforms by detecting and removing fraudulent contributions.
Tasks	• State of Art • The Problem • Competitive Analysis • The Need • The functional Analysis

	• Cost Analysis • MVP • Building our ML model • Build user interface

Detailed presentation of the project

Context

For some years now, the "phenomenon of the fake reviews" has increasingly crept into the marketing arena, capitalizing on rapid technological advances that enable the creation of artificial products for consumers. The concept of lies, imitation and deception has thus transcended its traditional boundaries, manifesting itself in a multitude of ways and permeating various aspects of our lives. The rise of fake reviews, strategically designed to influence opinions and reinforce reputations, further complicates the falsification network. These fake reviews, disguised as authentic testimonials, raise crucial questions about the integrity of online information and the reliability of digital spaces. In this context, the study of fake reviews has emerged as one of the key issues in digital marketing and social media research. Online product reviews, considered a form of electronic word-of mouth (eWOM), exert a significant influence on consumers' purchasing decisions.

Type of project

The "Detecting False Opinions on the Internet" project aims to develop an innovative and accessible solution for identifying false online opinions, based on the latest advances in artificial intelligence (AI) and natural language processing (NLP). The state of the art shows that, although solutions do exist, they are often costly and not very transparent. Our project addresses this problem by offering an efficient, transparent and affordable solution to the growing need for trust and integrity on review platforms. A competitive analysis reveals gaps in current solutions, underlining the need for a more accurate and user-friendly approach. Functional analysis of our solution includes review collection and analysis, detection of suspicious linguistic and behavioral patterns, and an intuitive user interface. Our MVP will include an AI-based detection model, a basic user interface, and filtering functionalities. The construction of the ML model will include data collection and preparation, model selection and training, as well as evaluation and optimization. Finally, the development of the user interface will integrate the latest web technologies to ensure an optimal user experience. All in all, this project is a comprehensive and innovative response to the challenges posed by fake reviews on the Internet, providing a reliable and transparent solution to this growing problem.

Budget forecast

For the "Detecting False Opinions on the Internet" student project, it's crucial to take all possible expenses into account, including human resources, technical infrastructure, software, and costs associated with promotion and training. Here's a detailed cost estimate:

1. Human resources

Personnel :

- Developers (4 students): €0, volunteer work as an academic project.
- Data Scientists (2 students): 0 €, volunteer work as an academic project.
- Supervisor/Encadrant: 0 €, included in the academic framework.

2. Technical infrastructure (if needed, otherwise €0)

Servers and Hosting :

- Development server: €100 per month x 12 months = €1,200
- Production server: €200 per month x 12 months = €2,400

3. Software and Tools

Licenses and subscriptions :

- Development environments (IDE, collaborative tools): €200 (annual licenses)
- Machine learning and data science tools (e.g. Jupyter Notebook, TensorFlow, etc.): Free, using open-source versions.
- Cloud services for AI model training (e.g. AWS, Google Cloud): €300 (free or reduced student credits)
- Data collection tools (web scraping): €100 (subscription to a scraping service).

5. Training and documentation

Training :

- Online courses and certifications ,Technical documentation and user guides: 0 €, created in-house by students

6. Promotion and Communication

Marketing and Promotion :

- Creation of project website: €100 (hosting and domain name)
- Online advertising and communication (social networks, blog posts): €200

7. Miscellaneous and Contingencies

Contingency reserve :

- Technical or logistical contingencies: €500

Budget Total Forecast

Categories	Cost Estimation (€)
Human resources	0
Technical infrastructure	3 600 (or 0)
Software and Tools	600 (or 0)
Data	700 (or 0)
Formation et Documentation	0
Promotion and Communication	300
Miscellaneaous and Contingencies	500
Total	**5700 (or 800)**

Justifications and potential adjustments

- **Technical infrastructure**: Costs can be optimized by using student credits offered by cloud services.
- **Software and tools**: Maximum use of open source tools to minimize costs.
- **Data**: Collaborate with companies to obtain data for free or at a reduced cost.
- **Training**: Seek scholarships or university funding to cover training costs.

The estimated budget for the "Detecting False Opinions on the Internet" project is €5700 or €800. This amount covers the essential aspects of the project, including technical infrastructure, software tools, necessary data, and promotion and training costs. By anticipating expenses and setting aside funds for contingencies, the team can manage resources efficiently and ensure the project's success.

Project organization

The 5 W method,

The 5 W method (CQQCOQP in French) method is a questioning technique used to structure and clarify a project by answering the questions: What, Who, When, How much, Where, How and Why. Here's how we applied this method to our project.

What (What is it?)

The "Detection of False Reviews on the Internet" project consists in developing a technological solution capable of identifying and filtering false reviews left on online platforms. The aim is to boost consumer confidence and protect honest businesses by ensuring that reviews reflect genuine experiences.

Who (Who's involved?)

- **Project team**: Developers, data scientists and analysts.
- **Stakeholders** :
    - **Consumers**: End-users of the solution.
    - **Companies**: Businesses and services concerned by reviews.
    - **Online Review Platforms**: Websites and applications hosting reviews.
    - **Supervisors and Teachers**: For academic supervision of the project

When (When should it be done?)

The project will be carried out in several phases with the following deadlines:

- **Phase 1:** Research and Planning (1st month)
- **Phase 2**: Data Collection and Preparation (2nd and 3rd months)
- **Phase 3**: Model Development (4th to 6th month)
- **Phase 4**: Application Development (7th to 9th month)
- **Phase 5**: Deployment and Monitoring (10th to 11th months)
- **Phase 6**: Communication and Promotion (12th month)

How much (How much does it cost?)

The estimated budget for the project covers :

- **Human resources**: Salaries for developers, data scientists and analysts.
- **Technical infrastructure**: Servers, databases, and development tools.
- **Software** : Licenses for data analysis and development software.
- **Training and Documentation**: Cost of writing guides and training users.
- **Promotion**: Advertising, conference participation and publications.

Where (Where is it taking place?)

- The project will be carried out mainly :
    - **On the university campus**: For research, development, and team meetings.
    - **Online**: Using collaboration and communication platforms for remote work.
    - **On online review platforms**: For data collection and deployment of the solution.

How (How is it done?)

The project will follow an agile project management methodology with the following steps:

- Research and analysis of existing methods.
- Collection and preparation of data from review platforms.
- Development of a model based on artificial intelligence to detect false reviews.
- Development of an application with an intuitive user interface.
- Deployment of the online solution and performance monitoring.
- Promotion of the project through various platforms and events.

Why (Why is this project important?)

Detecting fake reviews is crucial for :

- Strengthen consumer confidence by providing reliable and authentic reviews.
- Protect honest businesses from fraudulent practices that can damage their reputation.
- Improve the integrity of online review platforms, contributing to a more transparent and equitable digital ecosystem.

By answering these questions, we ensure a clear understanding and a well-defined structure for the "Detecting False Opinions on the Internet" project.

Project Management Tool

Project management tools are software solutions designed to help teams plan, execute, monitor, and close projects efficiently. These tools offer a range of features and functionalities to facilitate task management, scheduling, collaboration, communication, resource allocation, budgeting and reporting, enabling teams to streamline workflows, optimize productivity and achieve project goals and objectives effectively.

Example of project management tools we used for our project are Asana and GitHub.

Asana

We used Asana to enable us to create tasks, set priorities, assign responsibilities, and monitor progress through intuitive dashboards and timelines.

GitHub

GitHub is a web-based platform built around Git, a distributed version control system, that provides hosting for software development and version control using Git repositories. We used GitHub for this project in order to collaborate during development and track changes.

Constraints

A number of potential constraints may arise. Here is a detailed list of the main constraints the project team may face:

Technical constraints

1 - Data quality and quantity

- Difficulty obtaining a sufficient volume of notice data to train the model.
- Varied data quality, including unclear or poorly written notices.

2 - Algorithm complexity

- Development and optimization of complex artificial intelligence models.
- In-depth knowledge of natural language processing (NLP) and machine learning required.

3 - Infrastructures

- Limited computing resources (computing power, storage).
- Need for high-performance servers for data processing and application hosting.

Financial constraints

1 - Limited budget

- Limited funding for hardware and software resources.
- Difficulty allocating funds for unforeseen needs or project expansion.

2 - Development and maintenance costs

- Costs of acquiring quality data.
- Software licenses and development tools.

Human constraints

1 - Technical skills

- Varied level of skills and experience within the project team.
- Potential need for additional training in AI and software development.

2 - Team member availability

- Constraints related to students' academic schedules and other commitments.
- Coordination of efforts and management of schedules.

Organizational Constraints

1 - Project management

- Effective project management to meet deadlines and objectives.
- Coordination of different project phases to avoid delays.

2 - Communication and Collaboration

- Fluid communication between team members, especially if certain tasks are carried out remotely.
- Collaboration with external stakeholders, such as companies or opinion platforms.

Legal and ethical constraints

1 - Confidentiality and Data Protection

- Compliance with regulations on personal data protection (RGPD, for example).
- Anonymization of opinion data collected to avoid any invasion of privacy.

2 - Artificial Intelligence ethics

- Guarantee that the AI model is fair and free from discriminatory bias.
- Transparency on how the algorithm works and on the criteria for detecting false reviews.

Constraints on Use

1 - User acceptance

- End-user acceptance of the application and its results.
- User education on the importance of detecting false reviews and on the use of the application.

2 - User Experience (UX)

- Design of an intuitive, easy-to-use user interface.
- Feedback from users to continuously improve the application.

Competitive constraints

1 - Competition with other solutions

- Presence of similar solutions on the market.
- Need to differentiate the project through innovative features and improved accuracy.

By identifying and anticipating these constraints, the project team can develop strategies to overcome them and ensure the project's success.

Source pictures

https://technolense.com/bridging-the-digital-and-physical-world/