

**T.C. MALTEPE UNIVERSITY FACULTY OF ENGINEERING AND NATURAL SCIENCES DEPARTMENT OF SOFTWARE ENGINEERING**

**Software Project Management**

**TERM PROJECT**

**2024/25 Spring**

**SE 403**

**Business Requirements Document**

**AI-Generated Joke LLM Project**

**03/20/2025**

**Prepared By:**

|  |  |
| --- | --- |
| **Student Name** | **Student ID** |
| **Melih Küçük** | **21 07 06 048** |
| **Afnan Mohammed** | **21 07 06 811** |
| **Nurşeyda Doğan** | **21 07 06 047** |
| **İdil Öztürk** | **21 07 06 026** |
| **Hüseyin Eray Kızılkaya** | **22 07 06 302** |
| **Onur Ulaş Can** | **…** |

**Instructor: Ensar GUL**

[**ensargul@maltepe.edu.tr**](mailto:ensargul@maltepe.edu.tr)

**Table of Contents**

[**Software Project Management 1**](#_Toc193358819)

[**Business Requirements Document 1**](#_Toc193358820)

[**AI-Generated Joke LLM Project 1**](#_Toc193358821)

[**Table of Contents 3**](#_Toc193358822)

[**1. Business Requirements 4**](#_Toc193358823)

[**1.1 Project Objective 4**](#_Toc193358824)

[**1.2 Problem Statement 4**](#_Toc193358825)

[**1.3 Target Audience 4**](#_Toc193358826)

[**2. Background 4**](#_Toc193358827)

[**3. Business Opportunity 5**](#_Toc193358828)

[**4. Business Objectives 5**](#_Toc193358829)

[**5. Success Metrics 5**](#_Toc193358830)

[**6. Business Assumptions and Dependencies 5**](#_Toc193358831)

[**7. Scope and Limitations 5**](#_Toc193358832)

[**7.1 Major Features 5**](#_Toc193358833)

[**7.2 Limitations and Exclusions 6**](#_Toc193358834)

[**8. Business Context 6**](#_Toc193358835)

[**8.1 Stakeholder Profiles 6**](#_Toc193358836)

[**8.2 Project Priorities 6**](#_Toc193358837)

[**9. Deployment Considerations 6**](#_Toc193358838)

[**10. Functional Requirements 6**](#_Toc193358839)

[**11. Current Sprint Deliverables 6**](#_Toc193358840)

[**12. Next Sprint Plan 7**](#_Toc193358841)

[**13. Open Questions & Discussion Points 7**](#_Toc193358842)

# 1. Business Requirements

## 1.1 Project Objective

The objective of this project is to develop a **Large Language Model (LLM) capable of generating humorous and contextually relevant jokes in Turkish**. The model will take user inputs and generate witty, coherent, and culturally appropriate jokes.

## 1.2 Problem Statement

Existing language models often struggle with generating humor, particularly in Turkish. Furthermore, **most LLMs lack extensive support for the Turkish language, leading to unnatural or irrelevant joke outputs**. The goal of this project is to bridge that gap by creating an AI-powered joke generator tailored for Turkish-speaking users.

## 1.3 Target Audience

* **Individual Users** – For entertainment and humor
* **Educational Platforms** – For enhancing Turkish language learning through humor
* **Content Creators** – As a tool for generating comedic content

# 2. Background

Natural Language Processing (NLP) for Turkish is still evolving, and joke generation presents unique challenges in understanding context, punchlines, and humor structure. Given these challenges, we have chosen **Mistral** as our base model due to its **Turkish language support, ease of training, and strong performance in fine-tuning**.

# 3. Business Opportunity

* **High demand for AI-generated humorous content** in Turkish-speaking communities
* **Potential for viral AI-generated jokes on social media**
* **Innovative AI-driven humor for educational and entertainment applications**

# 4. Business Objectives

* **Develop a high-quality Turkish joke-generating LLM**
* **Ensure that generated content is contextually relevant and culturally appropriate**
* **Implement backend and frontend integrations for a seamless user experience**

# 5. Success Metrics

* **Evaluation of joke coherence and humor quality**
* **User feedback through surveys and A/B testing**
* **Accuracy in maintaining contextual integrity in joke generation**

# 6. Business Assumptions and Dependencies

* **Availability of sufficient Turkish joke datasets**
* **Adequate computational resources for fine-tuning the model**
* **Seamless backend and frontend integration for usability**

# 7. Scope and Limitations

## 7.1 Major Features

* **Mistral-based Turkish joke generation**
* **User-input-driven joke adaptation**
* **Web interface and API support**
* **Evaluation of database necessity for joke storage**

## 7.2 Limitations and Exclusions

* **No real-time content moderation yet**
* **Ethical and cultural sensitivity will be manually reviewed**
* **Humor perception varies; perfect joke generation cannot be guaranteed**

# 8. Business Context

## 8.1 Stakeholder Profiles

* **Development Team:** Responsible for training and integrating the model
* **End Users:** General audience seeking humor and entertainment
* **Academic Advisors:** Experts assessing the quality of generated jokes

## 8.2 Project Priorities

1. **Collection and fine-tuning of Turkish joke datasets**
2. **Integration of Mistral via Hugging Face**
3. **Development of backend and frontend infrastructure**
4. **Evaluation of database requirements for joke storage**

# 9. Deployment Considerations

* **Development Environment:** Google Colab & Jupyter Notebook
* **Deployment Strategy:** API-based accessibility
* **Infrastructure Requirements:** GPU-based computing resources for model inference

# 10. Functional Requirements

* **The model should generate coherent and contextually appropriate jokes.**
* **Response time should be within acceptable limits (<2 seconds).**
* **The model should be accessible through API integrations.**

# 11. Current Sprint Deliverables

* **Data Collection:** Gathering Turkish joke datasets
* **Data Cleaning:** Filtering offensive or inappropriate content
* **Data Structuring:** Formatting in JSON or CSV for training
* **Preliminary Model Testing:** Initial trials and performance evaluations

# 12. Next Sprint Plan

* **Optimize model performance and humor effectiveness**
* **Develop a web-based interface for public interaction**
* **Implement backend support for joke storage (if required)**
* **Evaluate frontend development feasibility using React**

# 13. Open Questions & Discussion Points

**To be discussed with the advisor:**

1. **Is it necessary to store generated jokes in a database?**
2. **If required, what would be the most suitable database choice?**
3. **Is database usage essential, or could caching suffice?**
4. **What are the best practices for implementing Mistral for this use case?**
5. **What additional considerations should be made when using Hugging Face models in production?**