Effort vs. Duration Report: Joke Generating LLM

Muharrem Şimşek - 210706001
Taha Demirhan - 210706002
Simay Aydın - 200706040
Aslımay Mısra Kandar – 210706043
Gökhan Mert Demirok - 210706311

08 April 2025

1. Project Overview

This document outlines the Effort vs. Duration analysis for the development of an LLM-based Joke Generation System. The system will generate jokes or short humorous stories based on user-provided prompts, utilizing a transformer-based machine learning architecture.

The project will be completed individually, so all Effort values reflect single-person work (person-hours).

2. Task Breakdown

Task No	Task Description	Effort (Person-Hourt)	Duration (Days)	Notes
1	Project Planning & Research	14	14	Research on transformers, LLMs, and joke generation techniques. Planning the project roadmap.
2	Dataset Collection & Preprocessing	14	7	Searching, cleaning, and formatting joke datasets suitable for model training.
3	Model Architecture Design	4	5	Designing the architecture of the transformer-based model for the specific task.
4	Environment Setup & Tooling	8	2	Setting up development environment, dependencies, and tools for model training.
5	Model Training & Tuning	30	28	Initial training, hyperparameter tuning, and handling overfitting or underfitting.
6	Prompt Engineering	10	3	Designing and testing different prompt structures to improve joke generation quality.
7	Evaluation & Testing	12	3	Testing the model's outputs, performing quality evaluation, and iterating improvements.
8	Final Optimization & Packaging	8	2	Final adjustments, cleaning code, and preparing the deliverable.
9	Report & Presentation Preparation	10	3	Writing project report and preparing the presentation slides.

3. Effort vs. Duration Summary

Total Effort	Total Duration
110 Person-Hours	67 Calendar Days

4. Additional Notes

• The project assumes 1-2 hours of work per day due to balancing with other coursework.

• Duration is affected by the computational time required for training large models.

• The scheduling allows for minor setbacks or additional tuning time.