Monthly Report

Project: Bangla Text Processing and Grammar Engine Development
Report Period: February

---

1. Introduction

This report summarizes the work completed in the previous month, outlining key achievements, challenges, and future objectives. The main focus has been on Bengali text extraction and processing, grammar engine development, and fundamental NLP implementations.

2. Descriptions with Figures

Bengali Text Extraction & Processing

- Successfully extracted Bengali text from images, PDF, excel, using Optical Character Recognition (OCR).
- Implemented text preprocessing steps such as: Removing punctuation, Filtering URLs, Eliminating noise, Cleaning dataset, Tokenization, Stemming, Morphological analysis recognition

Grammar Engine Development

- Developed a grammar-checking feature focusing on Present Tense detection.
- Trained a model to recognize four Present Tense forms accurately.
- Future plans include implementing Subject-Agreement rules and additional grammar rules.

NLP Core Progress

- Studied fundamental NLP concepts extensively.
- Applied NLP techniques in real-world projects to enhance practical knowledge.
- Continuous skill development through hands-on implementation.

Fresh One Project Progress

- Successfully completed assigned tasks for the project.
- Encountered initial errors in Bengali text processing, which were resolved systematically.
- Currently working on improving overall work quality.
- Aim to deliver even better outputs next month.

3. Code:

- [Github]
- [Kaggle]

4. Findings

- Bengali text extraction and processing require extensive cleaning and morphological analysis.
- Grammar model is capable of detecting Present Tense accurately, but further improvements are needed.
- Continuous project involvement has significantly improved NLP-related skills.

5. Discussion

- Addressed various challenges related to Bengali text preprocessing.
- Overcame errors in the Fresh One project with a structured debugging approach.
- Moving forward, advanced NLP techniques will be implemented for more robust models.

6. Work Completed

- Bengali OCR text extraction and preprocessing.
- Grammar engine Bangla Tense detection alhamdulillah comple.[Github]
- Advanced NLP concept exploration and project applications.
- Completion of multiple Fresh One project tasks. [kaggle]

7. Progress in Percentage

Task	Completion (%)
OCR	100%
Data Collection & Cleaning	90%
Grammar Engine (Bangla Tense Detection Engine)	90%
Core NLP Components	95%
Grammar Rule Engine (Rule-Based Layer)	85%
Linguistic Evaluation & Testing	80%
Comple One Project Tasks	80%

8. Future Task Targets

- Improve grammar engine by adding Subject-Agreement and other rules.
- Further refinement in Bengali text preprocessing.

- Work on expanding NLP implementation in real-world projects.
- Continue enhancing work quality in ongoing projects.

9. Conclusion
This month's progress has been significant in terms of Bengali text processing, grammar engine development, and NLP knowledge application. Moving forward, the focus will be on refining existing implementations, expanding project scope, and ensuring higher quality outputs in all tasks. The aim is to improve efficiency and effectiveness, delivering better results in the upcoming month.

Code Link:

Github link: https://github.com/bnlpe/Bangla-engine
Github link: https://github.com/bnlpe/Bangla-boss-engine