Byakaron

Bangla Grammar & Spell Checker

Nurul Huda
221902012

Jubaer Al Mamun
221902013

Md. Mursaline Parvez
221902363

Green University of Bangladesh

Supervised by
Dr. Muhammad Aminur Rahaman
Chairperson, Department of CSE

Table of contents

01 Introduction

02 Problem Statement

03 Motivation

04 Critical Analysis of the Related Works

05 Identifying Gaps

06 Research Question

07 Objectives

08 Proposed Methodology

09 Project Progress

10 Project Plan

11 Gantt Chart

12 Budget

13 Future Activity

14 Conclusion

15 References

16 Questions and Answers

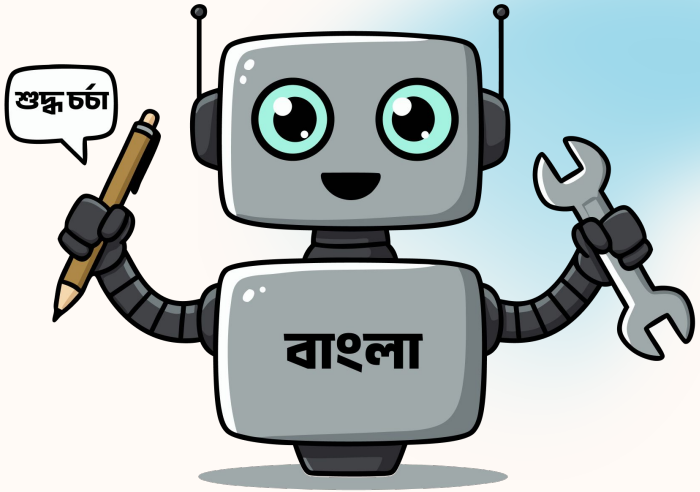
Byakaron

Bangla Grammar & Spell Checker

01

Introduction

1. Overview
2. Problem Statement



ব্যাকরণ বাবু

Overview

Definition and Purpose

Importance of Grammar Checking

Brief History of Bengali Language Processing

Features of Byakaron

- Grammar Checking
- Spell Checking
- Cross Platform

Problem Statement

Problem Statement

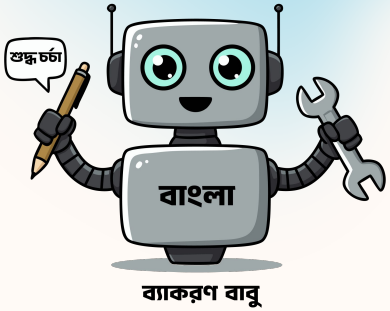
Byakaron

Bangla Grammar & Spell Checker

02

Motivation & Literature Review

1. **Motivation**
2. Critical Analysis of the Related Works
3. Identifying Gaps



Motivation

Enhancing Communication

Byakaron improves Bangla writing by enhancing grammar, spelling, and clarity.

Enriching Bangla

Byakaron aids Bangla learning with instant feedback, boosting user confidence, practice, and language proficiency.

Increasing Accessibility

Byakaron boosts access to Bangla grammar tools, helping users improve writing skills and confidence. And also being available on all major platforms.

Critical Analysis of the Related Works(1/2)

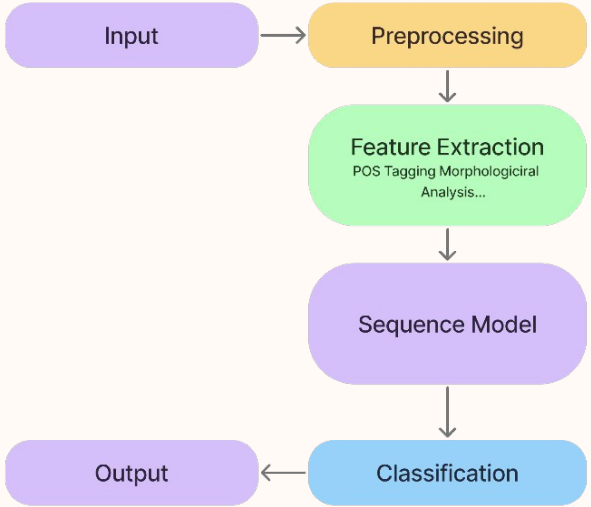


Figure 1: Workflow for rule-based or traditional method

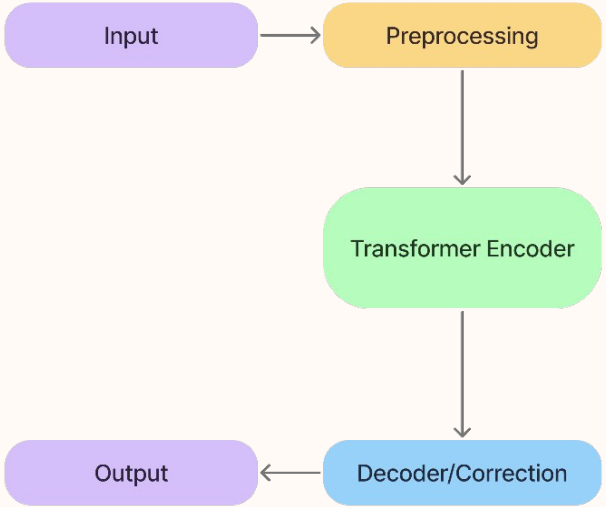


Figure 2: Transformer only pipeline

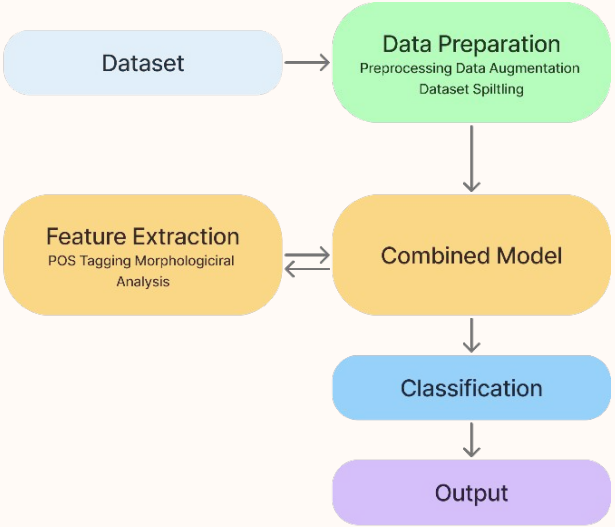


Figure 3: Workflow for hybrid neural network

Critical Analysis of the Related Works(2/2)

Table2: Some Research Papers Overview

Authors	Dataset	Contribution	Limitations
Bhattacharyya & Bhattacharya [3]	VAIYAKARANA Dataset	First standardized benchmark for Bangla grammar correction; evaluation framework	Limited coverage; lacks linguistic diversity; no real-time testing
Akter & Rahman [4]	Bangla Wikipedia + News Datasets	Comparative evaluation of BERT, T5, GPT-2 for Bangla grammar error detection	Does not address real-time performance; domain shifts not evaluated
Saha & Dey [5]	Custom Bangla Grammar Dataset	Fine-tuned BERT & T5 models showing improved accuracy in Bangla grammar correction	Limited dataset availability affects generalization
Islam & Kabir [6]	Large Bangla Text Dataset	Evaluation of BERT and T5 for sentence-level Bangla grammar correction	High computational cost; requires extensive resources for pre-training/fine-tuning

Identifying Gaps

Gaps in Current Research

01

Limited Real-Time Capability

02

Narrow Dataset Coverage

03

High Resource Demand

Opportunities for Improvement

01

Real-Time, Low-Latency Engine

02

Expanded, Diverse Dataset

03

Resource-Efficient Design

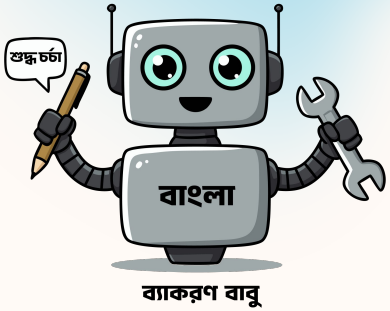
Byakaron

Bangla Grammar & Spell Checker

03

Research Question & Objectives

1. Research Question
2. Objectives



Research Question

Gaps in Current Research

Primary Research Question

What key features make Byakaron an effective Bangla grammar and spell checker, and how do they address unique linguistic challenges for Bangla speakers?

Secondary Questions

What key grammar issues do Bangla users face, and how can tools like Byakaron evolve with better features and user feedback?

Formulation of Research Question

Main Research Question

How does Byakaron enhance accuracy and UX while addressing Bangla grammar challenges?

Sub-Questions

How effective is Byakaron in correcting errors and learning from user feedback?

Relevance to Byakaron

How does Byakaron enhance Bangla grammar, spelling, and user engagement?

Objectives

ক খ গ ঘ ঙ
চ ছ জ ঝ ঞ
ট ঠ ড ঢ ণ
ত থ দ ধ ন

Specific Goals

- Accurate Bangla Spell Checking
- Grammar Error Detection
- User-Friendly Interface

Long-term Vision

- Contextual Suggestions and Style Improvements
- Bangla Language Ecosystem Integration
- AI-Powered Grammar Learning

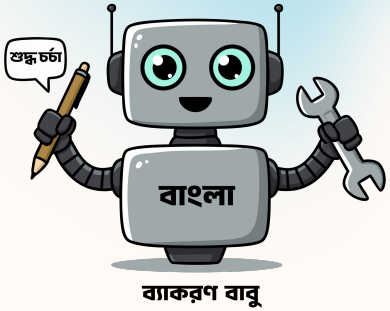
Byakaron

Bangla Grammar & Spell Checker

04

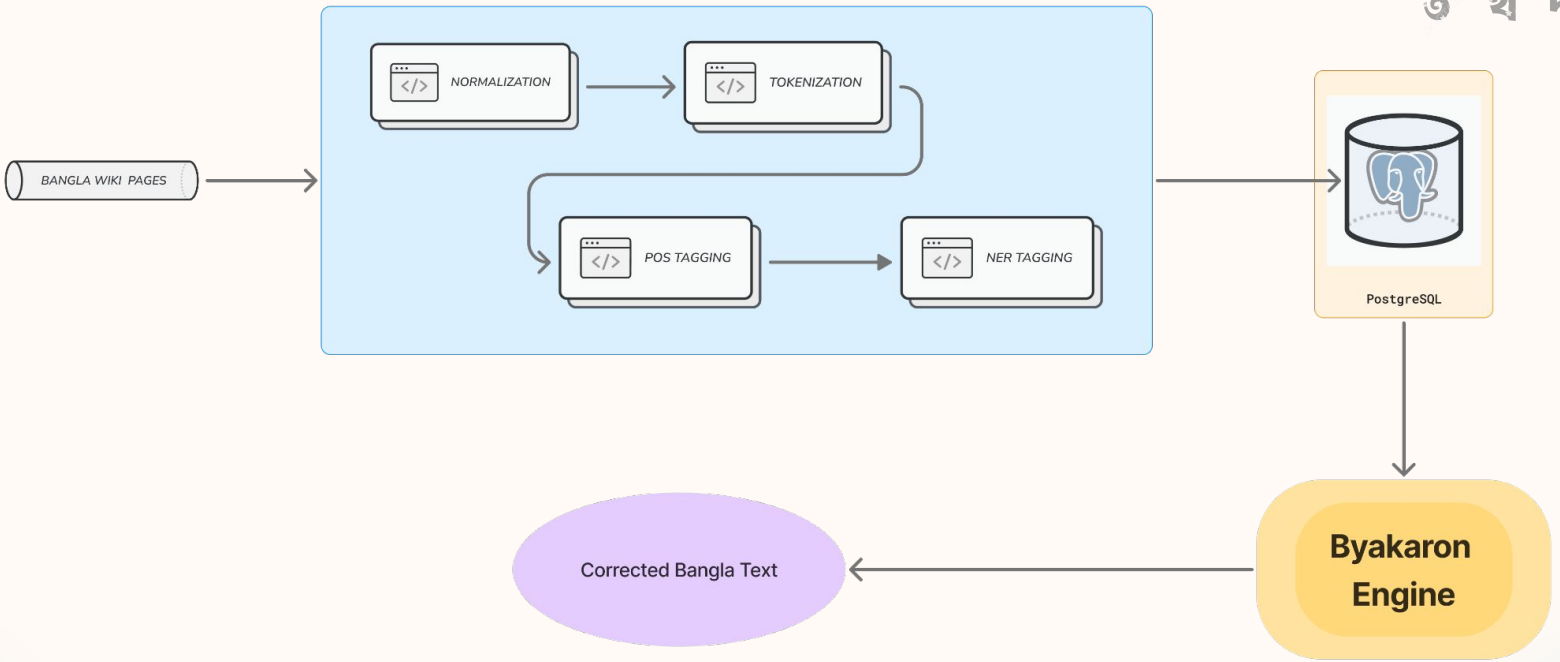
Methodology & Execution

1. Proposed Methodology
2. Project Progress



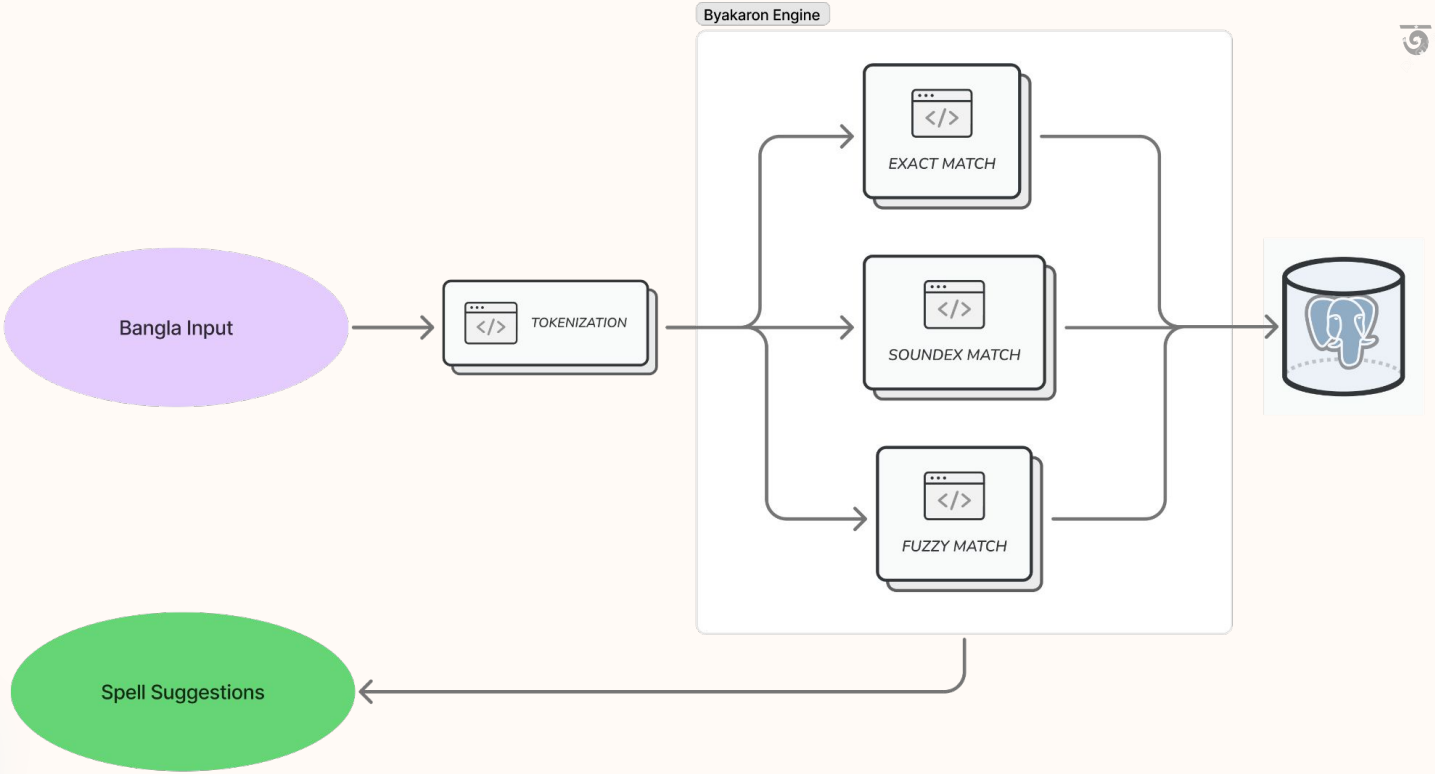
Proposed Methodology

Architecture



Proposed Methodology

Architecture



Proposed Methodology

Architecture

NORMALIZATION

<h1>আমি ভাত [2] খাই<h1>

↓

আমি ভাত খাই

TOKENIZATION

আমি ভাত খাই

↓

- আমি
- ভাত
- খাই

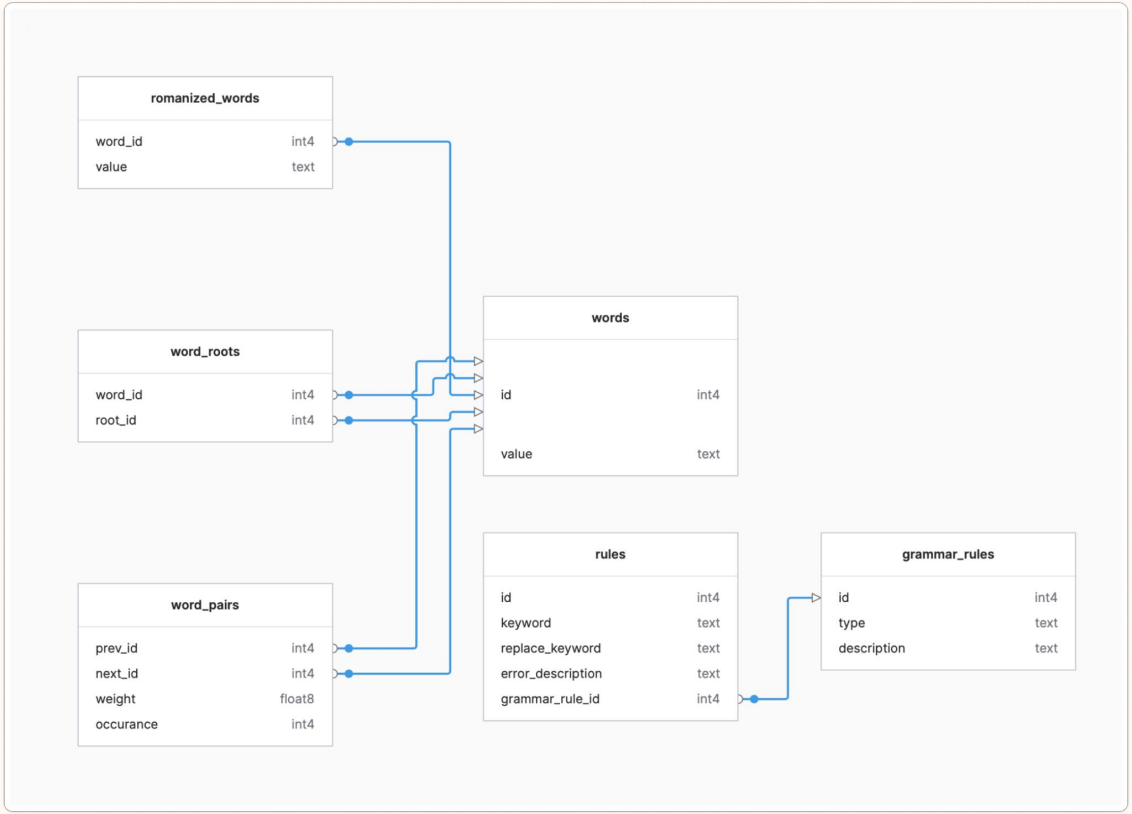
POS TAGGING

- আমি
- ভাত
- খাই

→

- আমি - SUB
- ভাত - OBJ
- খাই - VERB

Proposed Methodology (ER Diagram)



Function and Non-Functional Requirement

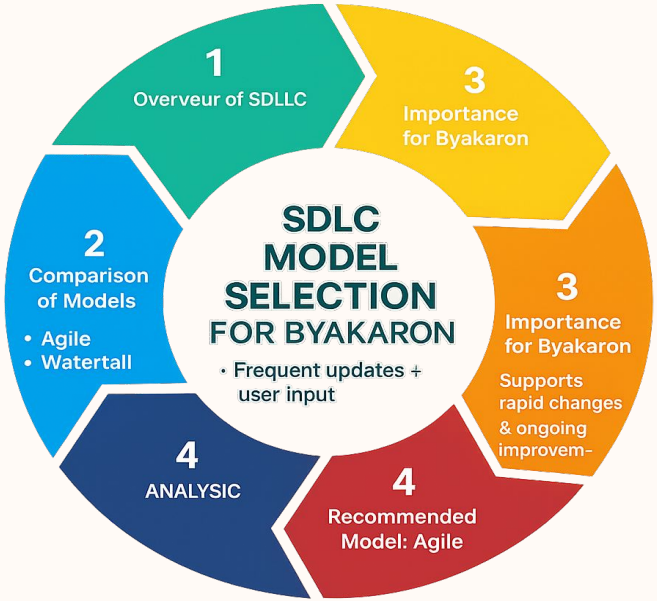
Functional Requirements

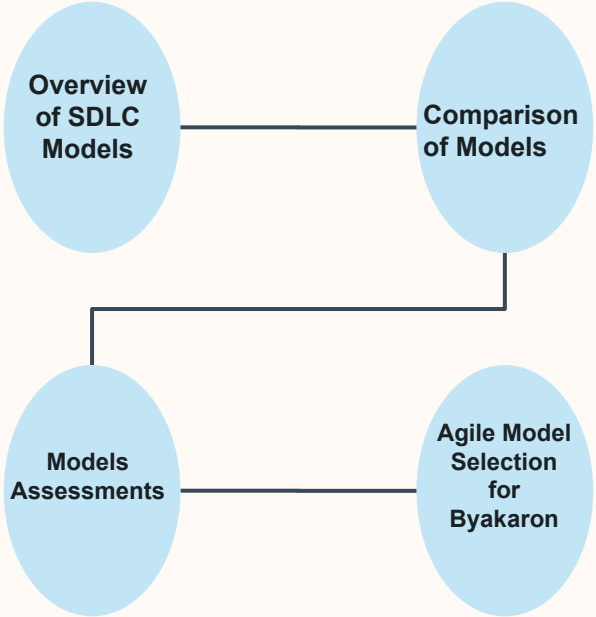
- Bangla Grammar & Spell Checking
- Real-time Suggestions
- Multiple Writing Modes
- Interactive Editing
- Custom Dictionary Support
- Export Corrected Text

Non-Functional Requirements

- Scalability
- High Performance
- High Availability
- User-Friendly Interface
- Security & Privacy
- Cross-Platform Compatibility

SDLC Model Selection





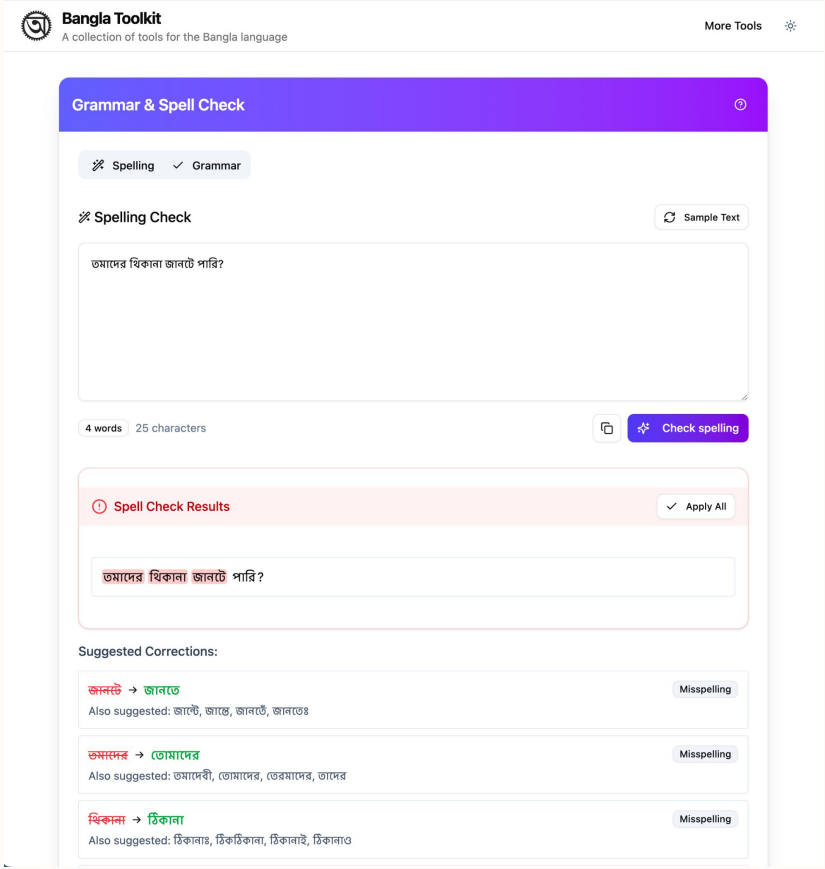
Project Progress (1/3)

Component	Status	Percentages
Dataset	Complete	100%
Engines	In Progress	50%
Grammar	In Progress	10%
Spell	In Progress	60%

Project Progress (2/3)

	Status			
Agenda	Dataset	Engines	Spell	Grammar
Research and Design	✓	✓	✓	✓
Development	!	✓	✓	!
Integration	✓	✓	✓	!
Final QA	!	!	!	!

Project Progress (3/3)



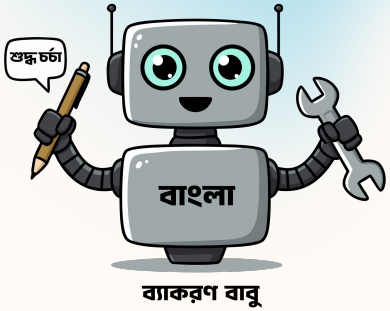
Byakaron

Bangla Grammar & Spell Checker

05

Planning & Monitoring

1. Project Plan
2. Gantt Chart
3. Budget



Project Plan

The risk plan covers tech issues, resources, and user adoption, with testing and feedback as key milestones for risk management.

Byakaron ensures timely, quality delivery through clear timelines, milestone tracking, and delay planning.

Byakaron follows three phases—R&D, testing, and launch—with milestones like prototype testing and user feedback for timely delivery.

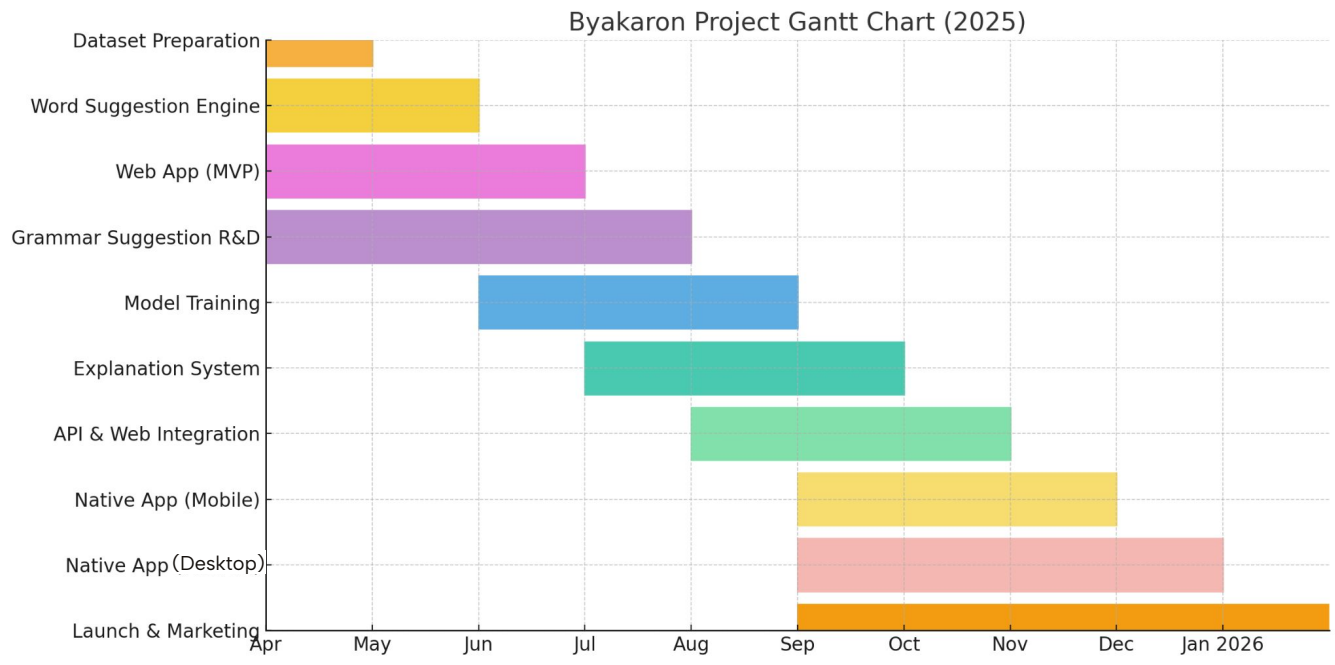
Risk Assessment

Deadline Management

Key Phases

শুদ্ধ ব্যাকরণ চর্চা

Gantt Chart


Byakaron Project Gantt Chart (2025)
Dataset Preparation
Word Suggestion Engine
Web App (MVP)
Grammar Suggestion R&D
Model Training
Explanation System
API & Web Integration
Native App (Mobile)
Native App (Desktop)
Launch & Marketing
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Jan 2026

Budget

Category	Description	Estimated Cost (BDT)
Developer Honorarium	Part-time support for ML, app, and web dev	1,50,000
Computing	GPU server rental/cloud compute for 3–4 months	80,000
Hosting & Backend Infra	Server, API hosting, domain, etc.	40,000
Mobile App Dev Tools	Store publishing, device testing	40,000
Marketing & Launch	Social media promotion, launch campaigns	30,000
Miscellaneous/Buffer	Contingency expenses	20,000
Total		**৳3,60,000**

শুদ্ধ ব্যাকরণ চর্চা

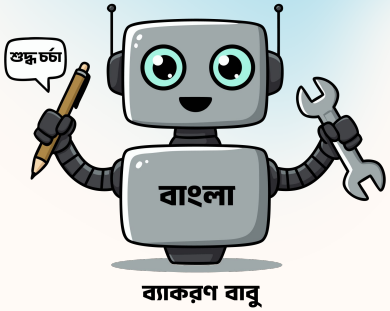
Byakaron

Bangla Grammar & Spell Checker

06

Future Activity & Conclusion

1. Future Activity
2. Conclusion
3. References
4. Questions and Answers



Future Activity

Short-Term Plans

01 — Upcoming Features

02 — User Engagement Strategies

03 — Feedback Collection

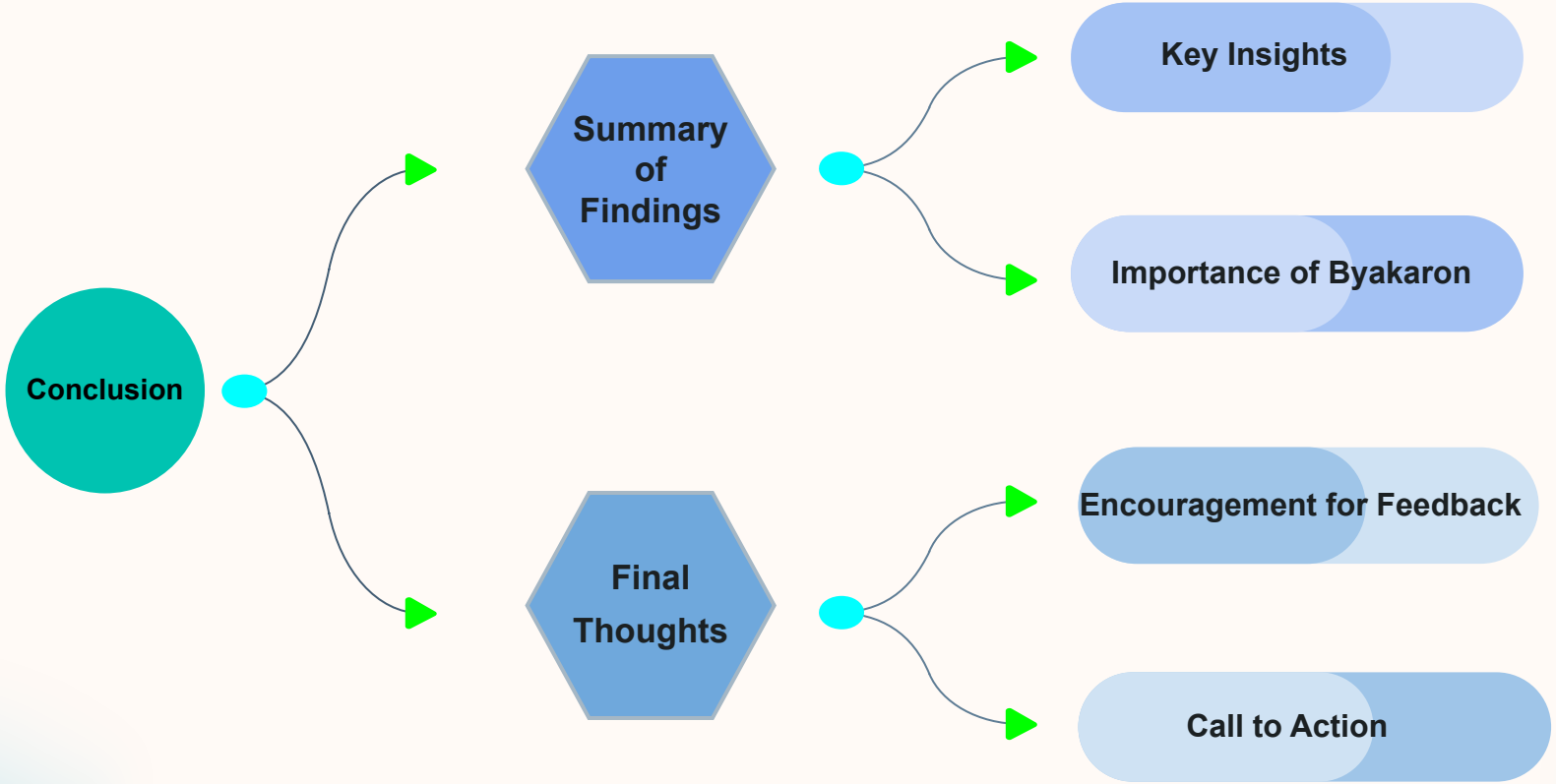
Long-Term Vision

03 — Sustaining Improvements

02 — Research Collaboration

01 — Expansion Plans

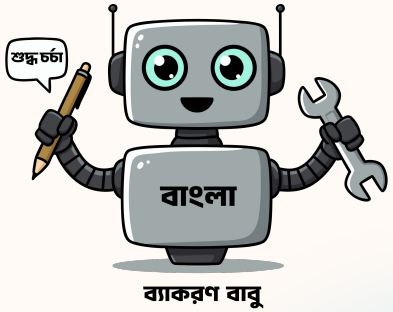
Conclusion


Conclusion
Summary of Findings
Key Insights
Importance of Byakaron
Final Thoughts
Encouragement for Feedback
Call to Action

References

1. https://dumps.wikimedia.org/bnwiki - Primary data source.
2. https://openbangla.github.io/ - Existing open source tools.
3. Bhattacharyya, P., & Bhattacharya, A. (2024). VAIYAKARANA: A Benchmark for Automatic Grammar Correction in Bangla. https://www.researchgate.net/publication/381580191_VAIYAKARANA_A_Benchmark_for_Automatic_Grammar_Correction_in_Bangla
4. Akter, A., & Rahman, M. (2023). Comparative Study of Transformer Models for Bangla Grammar Error Detection. https://arxiv.org/abs/2211.03730
5. Saha, S., & Dey, P. (2023). Bangla Grammar Error Correction Using Fine-tuned Language Models. https://www.academia.edu/9886588/NLG_Approach_for_Bangla_Grammatical_Error_Correction
6. Islam, R., & Kabir, M. (2023). Bangla Sentence Correction Using BERT and T5 Models. https://aclanthology.org/2023.banglalp-1.14.pdf

Byakaron

Bangla Grammar & Spell Checker

Thanks!