Intelligent Matchmaking for Diverse Needs Using NLP and Gradient Boosting for Inclusive Pairing

Ayana Lala, Bhumika Tekam, Jyoti Bhowmick and Archana Singh

ayanalala221061@acropolis.in, bhumikatekam220575@acropolis.in,

jyotibhowmick220071@acropolis.in, archanasingh220729@acropolis.in Dr. Shilpa

Bhalerao, Dr. Vandana Kate, Prof. Chanchal Bansal

hodcsit@acropolis.in,vandanakate@acropolis.in, chanchalbansal@acropolis.in

Department of Computer Science and Information Technology

Acropolis Institute of Technology and Research, Manglia, Indore-453771

Abstract:

Digital matchmaking platforms often fail to address the accessibility, emotional compatibility, and safety needs of individuals with disabilities. To bridge this gap, **this study proposes an AI-powered matrimonial platform** that integrates **Natural Language Processing (NLP)** and **Gradient Boosting Machine (GBM)** for intelligent and personalized matchmaking. **The proposed system** utilizes NLP to interpret emotional cues and user preferences, while GBM optimizes compatibility scoring to ensure matches are both data-driven and emotionally aligned. Key innovations include **accessibility features** (text-to-speech, voice command support) and a **Parental Vigilance module**, enabling guardians to monitor profile interactions—enhancing safety for users requiring assistance. **The framework** advances digital matchmaking by combining **AI-driven compatibility analysis** with **inclusive design principles**, fostering a secure and equitable experience for individuals with diverse needs. **This research** demonstrates how machine learning can transform matchmaking into a more empathetic, accessible, and trustworthy process for underserved populations.

1. Introduction

Finding a life partner is challenging, especially for individuals with disabilities, who often face barriers on mainstream matrimonial platforms. This research proposes an inclusive and secure matchmaking system tailored for differently-abled users, focusing on emotional compatibility and accessibility.

The platform uses **Natural Language Processing (NLP)** to understand user bios, preferences, and emotional tone, and **Gradient Boosting Machine (GBM)** to predict and improve match accuracy through continuous learning. These technologies form the foundation of intelligent, personalized matchmaking.

To ensure a safe and user-friendly experience, the solution includes features like **parental vigilance**, **voice assistance**, and **multi-choice inputs**, supporting users with mobility or
cognitive limitations. The following sections explain the system's architecture, matchmaking flow, and safety mechanisms in detail.

2. Literature Review

Online matrimonial platforms have transformed traditional matchmaking by allowing users to search for partners based on filters like age, caste, religion, education, and income [1]. In India, where marriage holds deep cultural value, these platforms are becoming widely accepted tools for partner selection.

Research such as *"Surfing for Spouses"* highlights how users balance traditional values with digital tools, showing interest in both family expectations and personal compatibility [2]. Other studies show that users prefer verified and well-written profiles, indicating the importance of trust and emotional compatibility in matchmaking [3].

Safety and accessibility remain major concerns, especially for women and people with disabilities. Studies recommend including features like identity verification and assistive technologies—such as screen readers, voice commands, and text-to-speech—for a more
inclusive experience [4][5]. Community support features like forums and group discussions can also boost user confidence, especially among disabled individuals [6].

Recent research suggests integrating technologies like **Natural Language Processing (NLP)** and **Gradient Boosting Machines (GBM)** to improve matchmaking. NLP helps analyze user-written bios and preferences using sentiment analysis to assess emotional compatibility [7][8]. NLP can also be adapted for accessibility with speech-to-text and voice navigation
features [9].

GBM is a machine learning model that combines small decision trees to predict compatibility scores from user data. It improves over time through learning from feedback and interaction
history. Ethical concerns such as fairness, explainability, and bias prevention are also emphasized in AI-driven matchmaking systems. Ensuring transparency in how matches are made is key to building user trust.

AI technologies are also being used to promote **social inclusion and digital support** for people with disabilities. In matchmaking, AI can help users connect more easily by understanding their needs and emotions. Research shows that platforms should do more than just meet basic accessibility—they should aim to **empower users** through features like adaptive interfaces and emotional understanding. This helps make matchmaking more personal and supportive, especially for those who face social or communication challenges.

Ethical use of AI is also very important in platforms made for sensitive groups. Studies highlight the need for **fair, transparent, and culturally aware** systems. Algorithms should not support
bias related to caste, gender, or disability. Newer methods like explainable AI and fairness-aware models help ensure that match suggestions are **just and inclusive**, supporting the goal of using technology for positive societal change.

In summary, existing platforms lack personalization and inclusivity. A system combining NLP and GBM with strong ethical and accessibility principles can offer a more empathetic and intelligent matchmaking experience.

3. Proposed Solution
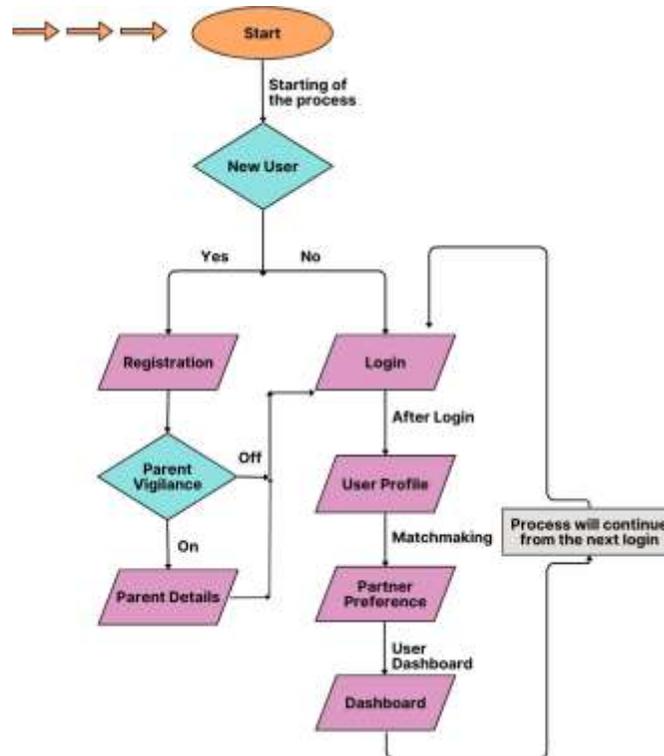
Fig.1. Flow diagram of the proposed solution

The diagram above outlines the key components and flow of the proposed inclusive matrimonial solution. Below is a summary of its working methodology.

3.1. System Flow Design

The solution follows a structured process that starts with user registration and goes through multiple stages. The methodology of the proposed solution includes the following steps:

- **Step 1: User Registration/Login**

  New users register, while returning users log in. If Parental Vigilance is enabled, parent contact details are collected for supervision.

Fig.2.(a) Registration page   **Fig.2.(b)** Login Page

- **Step 2: Profile Creation**

  Users complete their profiles using multiple-choice questions covering personal, lifestyle, and partner preferences.

- **Step 3: Partner Preference Collection**

  Users define desired partner traits (e.g., lifestyle, personality, disability needs) via structured questions.

- **Step 4: Matchmaking Process**

  Compatibility scores are calculated using **NLP and GBM**, considering:
- Shared interests
- Lifestyle alignment
- Disability-specific needs
- Emotional compatibility

- **Step 5: Dashboard Management**

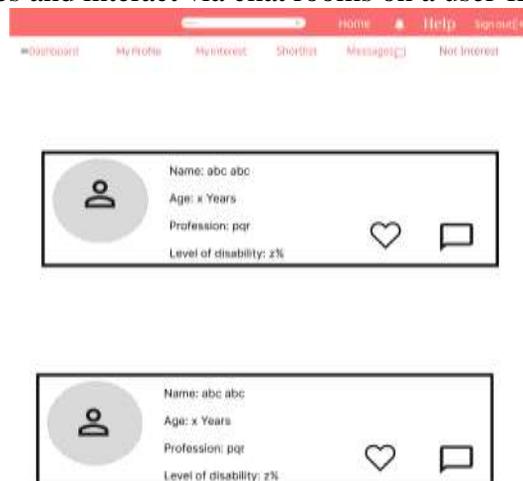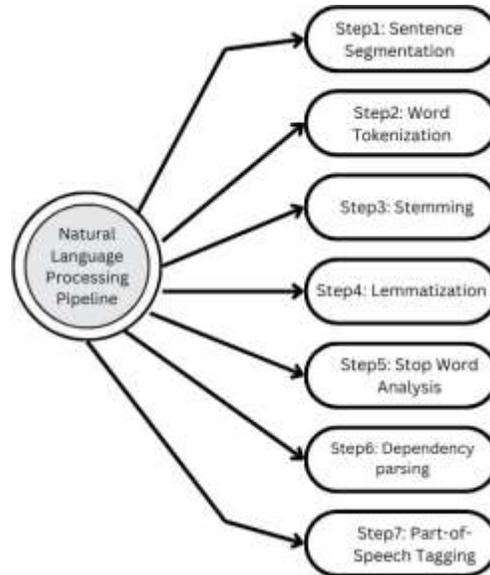 Users access matches and interact via chat rooms on a user-friendly dashboard.

Fig.3. User Dashboard showing matches

- **Step 6: Continuous Process**

  The system continuously refines suggestions using new data and user interactions.

3.2. NLP and GBM Implementation
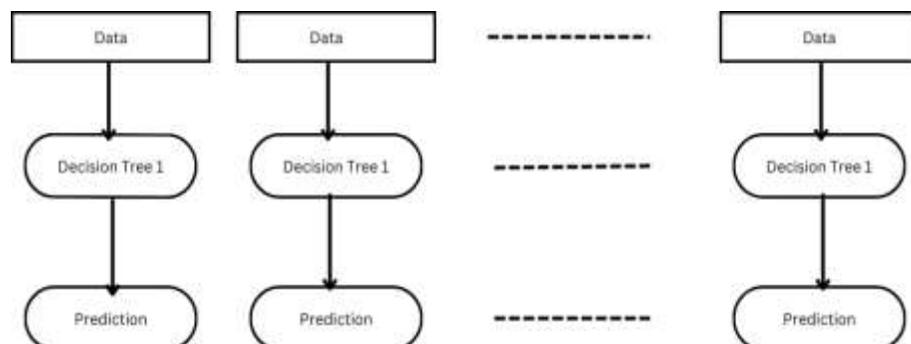
- **NLP for Data Analysis:**

Fig.4. NLP Algorithm

NLP processes user bios and preferences through:

- Sentence segmentation
- Tokenization, stemming, lemmatization
- Stop word removal
- Dependency parsing & POS tagging

Sentiment analysis evaluates emotional tone. **Pattern matching** identifies key traits for compatibility.

- **GBM for Compatibility Scoring:**

Fig.5. GBM Algorithm

The system uses XGBoost, an efficient GBM implementation:

- Trains decision trees sequentially
- Corrects errors iteratively
- Predicts refined compatibility scores
- Learns continuously from user feedback

3.3. Parental Vigilance Feature

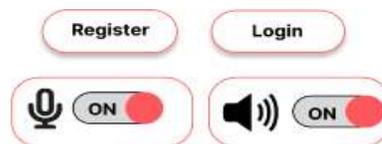When enabled, parents receive alerts on key activities like profile creation and chat interactions.
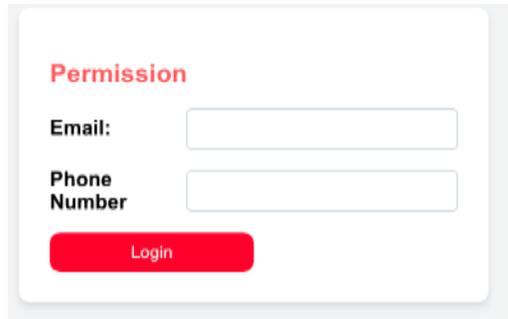
Fig.6. Parental Vigilance page of the proposed solution

3.4. User Accessibility Features

- The platform incorporates tools like:
    - **Text-to-Speech** for visually impaired users
    - **Voice Command** for users with limited mobility

Fig.7. Text to Speech and Voice Command Features

3.5. Data Handling and Privacy

- Ensure secure data encryption for user information.
- Implement two-step verification and profile verification via government ID or disability certification.

3.6. Evaluation Metrics

- The performance of the matchmaking algorithm can be evaluated using:
    - **Accuracy** of compatibility scores
    - **User Satisfaction** feedback
    - **Match Success Rate** (successful connections or relationships formed)

4. Result

This section evaluates the performance of the proposed compatibility prediction model, which uses **XGBoost** on user-preference data processed via **Natural Language Processing (NLP)**.

4.1 Prediction Accuracy – RMSE

The model achieved a **Root Mean Squared Error (RMSE) of 0.26**, indicating highly accurate predictions. This means the predicted scores deviate from actual values by only 0.26 units on average, which is minimal for a scale of 0 to 15.

4.2 Model Fit – R² Score

The model produced an **R² score of 0.98**, showing that 98% of the variation in actual compatibility scores is captured by the model, with only 2% attributed to unknown factors.

4.3 Visual Analysis

The scatter plot (Figure X) shows predicted vs. actual scores, with most points closely aligned along the ideal line—visually confirming strong model accuracy.
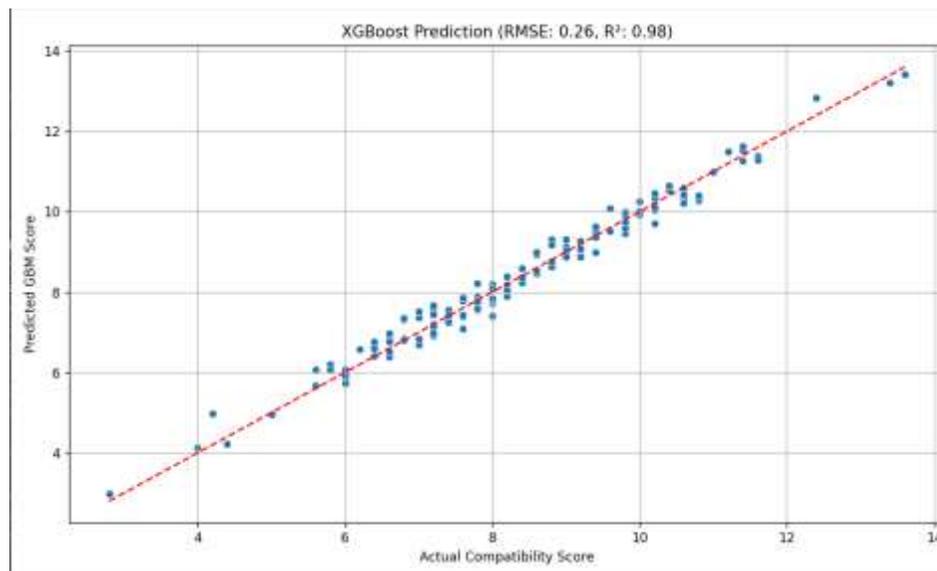
Fig.8. Graph showing relation between predicted and actual compatibility score

4.4. Summary of Findings

The low RMSE and high R² score confirm that the proposed **NLP + GBM** model is highly reliable and effective for predicting compatibility, especially on platforms designed for users with disabilities.

5. Conclusion

This research proposed an inclusive digital matchmaking solution for individuals with physical disabilities, powered by Natural Language Processing (NLP) and Gradient Boosting Machine (GBM) techniques. Designed with accessibility and user-centered features, the system prioritizes both ease of use and match accuracy.

By analyzing user preferences through NLP and predicting compatibility using XGBoost, the model achieved strong performance (RMSE: 0.26, R²: 0.98), showing high accuracy and reliability.

The proposed solution highlights how combining machine learning with empathetic design can enable smarter, more inclusive matchmaking. It empowers differently-abled users by offering personalized, data-driven match suggestions.

Future work could expand features like voice control, deeper emotional analysis, and adaptive learning for even better match precision. This lays the groundwork for ethical, intelligent, and accessible matchmaking in the digital era.

6. References

1. Raj, A., & Choudhary, H. (2020). Consumer Preferences Towards an Online Matrimonial Website.
2. Titzmann, F. (2011). Surfing for Spouses: Marriage Websites and the 'New' Indian Marriage?
3. Sharma, S., & Agarwal, P. (2019). Review on Matrimonial Information Systems and Services.
4. Jain, A. (2020). Matrimonial Websites and Changing Patterns of Marriage in India.
5. Saini, P., & Ahuja, N. (2022). Evaluating Quality of Matrimonial Websites: Balancing Emotions with Usability.
6. Das, D. (2018). Commercial Applications of NLP.
7. Kumar, V., & Yadav, A. (2020). Review on Natural Language Processing (NLP).
8. Singh, P., & Gupta, S. (2021). Commercial Use of NLP in Digital Platforms.
9. Kaur, A., Rani, R., & Arora, A. (2021). Uses of Natural Language Processing in Real-World Applications. Artificial Intelligence Review.