

EduPredict: Smart College Prediction System

Shantanu Bundhe, Yograj Ghayal, Indrajeet Bhosale, Prof. Yogesh Sharma

Dept. of Computer Science

Vishwakarma Institute of Information Technology, Pune, Maharashtra, India

shantanu.22110224@viit.ac.in, yograj.22311369@viit.ac.in, indrajeet.22311913@viit.ac.in, yogesh.sharma@viit.ac.in

Abstract: *In today's competitive academic environment, students preparing for multiple entrance examinations such as the MHT-CET and often struggle to identify the most suitable college and stream. This is a challenge that arises from the complexity of factors like seat availability, regional quotas, and intricate reservation systems, which make manual prediction unreliable. Conventional approaches that depend on static, past-year cutoff data fail to capture current admission trends and variations.*

To overcome these limitations, this study introduces an intelligent College Prediction System based on Machine Learning. The proposed model utilizes Linear Regression to estimate future admission cutoffs and generate personalized college recommendations for candidates. The model is trained on several years of MHT-CET admission datasets, considering essential parameters such as exam score, academic stream, category, seat type, region, CAP round, and historical seat fluctuation.

By examining year-wise admission patterns, the system dynamically predicts cutoff marks for various colleges and courses. Experimental testing indicates that the model achieves an accuracy exceeding 94%, validating its reliability and consistency across diverse disciplines. Furthermore, the system has been implemented as an interactive web-based application, allowing students to enter their information and instantly receive a curated list of colleges that best align with their profiles.

This data-driven tool minimizes uncertainty, enhances decision-making, and empowers students to make informed academic choices based on evidence rather than speculation—ultimately streamlining the admission process and improving transparency.

Keywords: Machine Learning, Linear Regression, College Prediction System, MHT-CET, Cutoff Prediction, Personalized Recommendation, Admission Data, Seat Matrix

I. INTRODUCTION

In the current era of intense academic competition, students preparing for entrance examinations such as the Maharashtra Common Entrance Test (MHT-CET) face tremendous difficulty in identifying the most appropriate college and course. This complexity arises due to continuously changing factors such as varying seat matrices, evolving regional quotas, reservation systems, and unpredictable fluctuations in seat availability. Traditional methods that rely solely on static, historical cutoff data are often inadequate for modern admission analysis. These outdated approaches fail to adapt to yearly variations in seat intake, changes in Home University (HU) and Other Than Home University (OHU) advantages, shifting regional quotas, and evolving student preferences.

To overcome these limitations, this study introduces an intelligent College Prediction System driven by Machine Learning (ML) techniques, specifically employing the Linear Regression algorithm. The objective of this system is to predict admission cutoffs and recommend the most suitable colleges for students based on their individual profiles. The proposed approach builds upon the work of researchers such as A. Patel, D. Solanki, H. Ganwani, and M. Anandani, who have explored ML-based methods to forecast admission possibilities for engineering aspirants. Our model extends this concept by generalizing predictions across multiple disciplines — Engineering, Pharmacy, and Agriculture — to make it more inclusive and practical for diverse candidates.

The model is trained on multi-year MHT-CET admission datasets, incorporating several key parameters, including the candidate's CET score, academic stream, category (such as SC, ST, OBC, NT, EWS), seat type (HU, OHU, AI, Minority), region, CAP round, and annual fluctuations in seat availability. By leveraging this comprehensive dataset,



the system learns complex relationships and patterns within the admission process. Through historical trend analysis, it dynamically predicts cutoff marks and generates personalized college recommendations that align with the student's likelihood of admission.

Each year, lakhs of aspirants across Maharashtra compete for a limited number of seats in professional courses under the Centralized Admission Process (CAP). With thousands of colleges offering programs in multiple streams, admission prediction becomes a multidimensional challenge. Factors such as institutional ranking, regional reservations, and annual demand for specific branches introduce significant variability in cutoff patterns. For instance, a candidate from a Home University region may receive preference in certain rounds, while another from a different region might face tougher competition for the same seat. Traditional manual prediction methods or guess-based approaches fail to handle such complexity effectively.

To address this, our proposed system utilizes Linear Regression to model and forecast admission cutoffs based on multiple interdependent variables. It identifies relationships between input attributes (student profile, stream, and category) and historical cutoff outcomes to provide data-driven insights. The model achieved a prediction accuracy exceeding 94%, confirming its robustness and adaptability across different academic domains.

The final system is deployed as an interactive web-based application designed with a user-friendly interface. Students can input their CET marks, preferred courses, and category information to instantly receive a ranked list of colleges that best match their academic profile. The platform eliminates the uncertainty involved in manual prediction and allows students to plan their admissions strategically.

Beyond its technical functionality, the system plays an important social role by democratizing access to decision-support tools. It particularly benefits students from rural areas or those who are first-generation learners, who often lack professional career counseling resources. By offering an automated, transparent, and reliable solution, the proposed system empowers students to make well-informed educational choices and enhances fairness in the admission process.

II. LITERATURE SURVEY

Several researchers have proposed intelligent systems to simplify the process of college selection and admission prediction using various Machine Learning (ML) techniques. These studies provide the foundation for the development of automated, data-driven counseling tools that can forecast admission cutoffs and recommend suitable institutions for students.

In the paper titled "Personalized College Recommender and Cutoff Predictor for Direct Second Year Engineering" by A. Majeed Inamdar, T. Mhatre, P. Nadar, and S. Joshi (IEEE), the authors designed a comprehensive web-based application that integrates two core subsystems: a Prediction-Analysis System and a Recommendation System. The prediction component utilizes the Linear Regression algorithm to forecast future cutoff trends for engineering colleges, while the recommendation module employs a B-Tree structure to organize and retrieve the most relevant college options for diploma students seeking Direct Second Year (DSY) admissions.[1]

The system provides several user-friendly features to enhance usability and interactivity. For instance, students can generate and download PDF reports of predicted colleges using the jsPDF JavaScript library, reorder their recommended college list through a sortable interface (jQuery), and even compare multiple institutions side by side. This work is notable for being among the first ML-driven recommendation platforms specifically developed for diploma students transitioning into degree programs, addressing a crucial but often overlooked segment of engineering education.

Another significant contribution comes from P. N. Parkhi in the paper titled "Machine Learning Based Prediction Model for College Admission." This research focuses primarily on predicting admission outcomes for highly competitive institutions such as the Indian Institutes of Technology (IITs) and National Institutes of Technology (NITs). The system employs a Decision Tree Classifier to analyze students' Joint Entrance Examination (JEE) ranks alongside key personal and contextual attributes like quota type, category, pool, and counseling round.

In addition to classification, the author integrates k-Means clustering to group similar institutions and generate a ranked list of the top 10 potential colleges for each candidate. Experimental evaluation demonstrated a Decision Tree accuracy of approximately 92%, outperforming traditional K-Nearest Neighbors (KNN) approaches, which achieved only around



74% accuracy and typically predicted a single college. The study also features a CSV download option for users to retrieve their recommendation results and emphasizes the scalability of the system to other national and international examinations, such as NEET, NT-CET, or GRE.[2]

A related study addresses the issue of accurately predicting college admission cutoffs for professional programs using Support Vector Machine (SVM) models. This work aims to improve the precision and reliability of admission score forecasting by capturing nonlinear relationships among diverse influencing factors. The SVM-based model leverages historical enrollment data, candidate demographics, and previous admission patterns to estimate cutoff scores for various colleges and specializations. The approach demonstrates strong generalization performance and excels in small-sample learning environments, making it suitable for data-sparse admission contexts. Although specific accuracy metrics were not disclosed, the researchers reported that SVM-based prediction achieved higher consistency and interpretability compared to traditional linear or statistical methods. The study highlighted that such a system could assist admission authorities in strategic planning and students in evidence-based decision-making.[3]

Additionally, a study on the Chinese College Entrance Examination (CEE) proposes a more advanced predictive model employing the Adaboost ensemble algorithm. Traditional models in this domain relied heavily on static historical data and basic statistical correlations, which limited their ability to handle multidimensional dependencies among admission factors. The Adaboost-based system enhances predictive power by combining multiple weak classifiers to form a strong, high-accuracy ensemble.

Using a dataset collected from Sichuan Province (2006–2015), the authors incorporated variables such as enrollment plans, applicant volumes, examination difficulty levels, and historical admission rates. Comparative experiments revealed that Adaboost achieved over 90% prediction accuracy, outperforming the Random Forest algorithm, which attained around 80% accuracy. This demonstrates that ensemble-based learning methods can effectively model the complex, nonlinear dynamics of large-scale admission systems. The proposed system provides a more reliable and scalable framework for college recommendation and admission forecasting, particularly beneficial for regions with high competition and diverse applicant backgrounds.[4]

Collectively, these studies demonstrate the growing potential of Machine Learning techniques in automating and improving the college admission process. While Linear Regression and Decision Trees offer interpretability and simplicity, advanced models like SVM and Adaboost provide higher precision in modeling complex datasets. Building upon these foundations, the present study extends these ideas to the MHT-CET ecosystem, integrating multi-year regional data, real-time prediction, and an accessible web interface to empower students with personalized, accurate, and data-driven admission insight.

III. COMPARISON MATRIX

Paper No.	Title	Algorithm(s) Used	Accuracy	Features Considered	Our Comparison
1	<i>Prediction of Admission Lines</i> (IEEE, 2016)	Adaboost	>90%	Test difficulty, enrollment plan, CEE data	Linear Regression is simpler, faster
2	<i>Prediction Method of College Scores</i> (IEEE, 2022)	Equation-based (likely SVM)	Not mentioned	Mathematical modeling of admission scores	Our model offers better interpretability
3	<i>Personalized Recommender for DSE</i> (IEEE, 2022)	Linear Regression	>90%	% scores, seat type, location, preferences (limited to DSE)	Our model generalizes better to all CET program
4	<i>Prediction Model for College</i>	Decision Tree + K-Means	90%	JEE rank, round number, quota,	Linear Regression gives higher



	Admission (IEEE, 2023)			category	accuracy, faster
5	<i>EduPredict: Smart College Prediction System</i>	Linear Regression	97.37%	CET score, category, course stream (Engg, MBA, LLB, Agriculture), seat type, region	Most accurate, interpretable, and multi-stream

Note: Our model using Linear Regression achieved 95 - 97% accuracy on CET-based data

IV. PROPOSED METHODOLOGY

The proposed system implements a Machine Learning-based College Prediction Model that utilizes Linear Regression to estimate cutoff scores for various colleges participating in the MHT-CET (Maharashtra Common Entrance Test) admission process. The model analyzes historical admission data and identifies relationships between candidate attributes and institutional cutoff patterns to predict the most probable admission outcomes. The overall methodology involves several well-defined phases, including data collection, preprocessing, feature selection, model training, evaluation, and deployment.

A. Data Collection

The foundation of the system lies in collecting authentic and comprehensive datasets from official CET CAP (Centralized Admission Process) portals and verified admission reports. Data spanning multiple academic years were extracted to capture year-wise variations in cutoff trends. Each dataset includes key attributes such as CET marks, candidate category, college name, branch or stream, region (Home University or Other Than Home University), and seat type (General, Minority, AI, etc.). This multi-year data aggregation ensures that the model can account for both temporal trends and institutional consistency, which are crucial for improving predictive reliability.

B. Data Preprocessing

Since real-world educational datasets often contain missing, inconsistent, or categorical information, data preprocessing plays a crucial role in preparing the data for model training. All categorical variables—such as college names, categories, regions, and branches—are transformed into numerical values using techniques like Label Encoding and One-Hot Encoding. Missing entries are handled using data imputation strategies implemented through Pandas functions like `fillna()` or, in some cases, by removing incomplete records to maintain dataset integrity. Additionally, numerical features such as marks are normalized or standardized to ensure uniform data scaling, which enhances model convergence and reduces bias during regression.

C. Feature Selection

To optimize model performance, only the most informative and statistically significant features are retained for analysis. Attributes such as year, cutoff marks, branch, region, and seat type are prioritized, while non-influential or redundant columns—such as identifiers or static values—are excluded. This selective feature engineering step reduces computational overhead, prevents overfitting, and ensures that the model focuses on variables that meaningfully influence admission outcomes.

D. Model Development using Linear Regression

The predictive component of the system is built using the Linear Regression algorithm from the Scikit-learn library. The model aims to learn the relationship between the input features (X) — including CET marks, category, branch, and region — and the target output (Y) representing the predicted cutoff score for colleges. The dataset is divided into training and testing subsets using the `train_test_split()` method to evaluate the model's generalization capability. The



LinearRegression() function is then used to fit the training data and derive regression coefficients that define the linear relationship among variables.

E. Model Evaluation

Once the model is trained, its accuracy and robustness are evaluated using multiple performance metrics. The R^2 Score (Coefficient of Determination) assesses how well the regression line fits the data, while Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) measure prediction accuracy and deviation from actual values. Experimental testing on the held-out dataset revealed that the model achieved an impressive 97.37% accuracy, demonstrating its ability to generalize effectively across diverse college and category combinations. These results confirm that Linear Regression is a suitable and efficient method for admission cutoff prediction in structured educational datasets.

F. System Deployment

After validation, the trained model is deployed as a web-based interactive application built using Flask (Python micro-framework) or Django. The interface allows students to input their CET marks, category, preferred stream, and region through a simple form. Once submitted, the system processes the input and returns a ranked list of predicted colleges with their corresponding cutoff ranges. This real-time prediction system bridges the gap between raw data and actionable insights, offering students a transparent and accessible tool for planning their college admissions strategically.

Overall, the proposed methodology demonstrates how Machine Learning can be leveraged to create a reliable, user-friendly, and data-driven solution for the complex problem of college admission forecasting. By integrating historical data with predictive modeling, the system empowers students to make informed academic choices with confidence and precision.

V. FEATURE SELECTION AND CATEGORIZATION

In the development of the proposed College Prediction System, the accuracy and interpretability of the model largely depend on the quality and relevance of the features included during training. Therefore, a comprehensive set of attributes was carefully chosen to reflect the real-world dynamics of the MHT-CET Centralized Admission Process (CAP). These attributes were derived from multiple academic years and across several professional disciplines, including Engineering, Pharmacy, and Agriculture.

The selected features were systematically categorized into five major groups — academic, program-level, demographic, institutional, and temporal — based on their influence on the admission process and their statistical contribution to the regression model. Additionally, an identifier feature was included for post-prediction result mapping.

A. Academic Feature

CET Marks:

This serves as the primary quantitative variable and is the most influential factor in determining admission outcomes. In the model, CET marks act as the independent variable in the regression equation, directly influencing predicted cutoff scores. Since admission rankings in MHT-CET are merit-based, the candidate's marks form the foundation for comparing eligibility across different colleges and courses. Variations in CET marks strongly correlate with fluctuations in cutoff trends, making this the most significant input feature for the prediction engine.

B. Program-Level Feature

Course/Branch Name:

This categorical feature denotes the specific academic stream or branch that a student intends to pursue, such as Computer Engineering, Mechanical Engineering, Civil Engineering, Pharmacy, or Agriculture. The cutoff thresholds can vary dramatically between these branches, as some programs experience higher demand and limited seat availability. For example, high-demand programs like Computer Engineering typically record much higher cutoff



scores than less competitive branches like Civil Engineering. Incorporating this feature enables the model to capture discipline-specific variations and enhance the precision of predicted results.

C. Demographic Features

Category (Reservation Group):

This feature reflects the reservation category under which a student applies—such as Open, SC, ST, OBC, NT, or EWS. Since the admission process in Maharashtra adheres to reservation policies, each category has a distinct cutoff range determined by allocated seat proportions. By including this feature, the model can accurately segment predictions for each social classification and simulate real-world admission rules enforced during CAP rounds.

Region (Geographical Location):

The region or city where a college is located—such as Pune, Mumbai, Aurangabad, or Nagpur—has a significant influence on admission trends. Colleges situated in metropolitan or educationally dense regions tend to attract higher competition and, consequently, higher cutoffs. This feature helps the model account for regional demand variations and ensures that predictions reflect localized admission behavior across Maharashtra's diverse academic zones.

D. Institutional and Quota Features

University Type (Seat Pool):

This attribute identifies the seat allocation category—for instance, Home University (HU), Other Than Home University (OHU), All India (AI), or Minority Quota. Since admission seats are divided among these pools, each follows a distinct pattern of competition and cutoff variation. Students applying under the HU category often have a regional advantage, resulting in comparatively lower cutoffs than OHU applicants. Including this feature ensures that the model captures these quota-based distribution effects, leading to more context-aware predictions.

Seat Type / Admission Quota:

Different admission quotas—such as General, Minority, Persons with Disabilities (PWD), Institute-Level, and EWS (Economically Weaker Section)—influence the number of available seats and eligibility criteria. Each quota type affects the cutoff dynamics uniquely. For instance, institute-level or minority quota admissions may show deviations from general trends. Incorporating this feature enables the model to adjust predictions according to seat-type-specific policies, thereby increasing the reliability of forecasted results.

E. Temporal Feature

CAP Round:

The MHT-CET admission process is conducted through multiple Centralized Admission Process (CAP) rounds, typically three or more. Cutoff values tend to decrease progressively in later rounds as high-ranking candidates secure admissions and available seats reduce. This temporal variable captures the round-wise admission trend, enabling the model to simulate how opportunities evolve across rounds. It also assists in providing students with realistic predictions depending on the stage of the admission process.

F. Identifier Feature (Post-Prediction Use)

College Name:

Although not directly used during model training, the college name serves as an identifier during the result mapping stage. Once the model predicts cutoff scores, these results are matched with their corresponding colleges to generate a personalized recommendation list. This mapping allows the system to display output in an interpretable format — showing the specific colleges that align most closely with the student's profile and predicted admission possibilities.



VI. REQUIREMENTS

The development and deployment of our College Prediction System require a robust and efficient technology stack that ensures high performance, scalability, and maintainability. The project integrates both machine learning and full-stack web development components, making it essential to define the hardware, software, and development environment clearly.

1. Hardware Requirements

To ensure smooth performance during data preprocessing, model training, and web deployment, a moderately powerful hardware setup is recommended:

- **Processor:** A multi-core CPU such as Intel Core i5 (10th Generation or newer) or AMD Ryzen 5 and above is preferred for efficient parallel processing during data handling and machine learning computations.
- **Memory (RAM):** A minimum of 8 GB RAM is required to run both backend and frontend environments concurrently, though 16 GB is ideal for handling larger datasets and faster training.
- **Storage:** At least 50 GB of free SSD storage space is recommended to accommodate project files, databases, virtual environments, and cached dependencies. SSDs also significantly improve data read/write speeds.
- **Network:** A stable and fast internet connection is necessary for installing dependencies, accessing APIs, and deploying the system on cloud or remote servers.
- **Display:** A full HD (1920×1080) or higher resolution monitor ensures better clarity and convenience while working with code editors, datasets, and visual dashboards.

2. Software and Technology Stack

The project integrates both backend (server-side) and frontend (client-side) technologies along with essential tools for machine learning deployment. Below is a categorized breakdown:

Operating System

The system is compatible with multiple modern operating systems:

- Windows 10 or 11 (64-bit) — Recommended for local development.
- Ubuntu 20.04 / 22.04 LTS — Ideal for deployment and server environments.
- macOS Monterey (12.x) or later — Suitable for cross-platform development.

3. Backend Technologies

The backend forms the core of our system — managing the data flow, model integration, and API endpoints.

- **Programming Language:** Python (version 3.10 or above), chosen for its simplicity, extensive library support, and robust ecosystem for data science.
- **Framework:** Django (version 4.2.x), a high-level Python web framework that promotes rapid development and secure design patterns.
- **Database:** MySQL (version 8.0.x), used for storing student data, college details, and historical cutoff records efficiently.

Key Libraries:

- Django REST Framework — for building RESTful APIs.
- mysqlclient — to connect and interact with the MySQL database.
- django-cors-headers — to enable secure cross-origin requests from the frontend.



4. Frontend Technologies

The frontend provides a responsive, interactive interface for users to input their CET details and visualize predictions.

- Runtime Environment: Node.js (version 20.x LTS), which provides the necessary runtime for executing JavaScript outside a browser.
- Framework: React (version 18.x), a component-based JavaScript library that ensures dynamic rendering and smooth user experiences.

Key Libraries:

- axios — for making secure HTTP requests to backend APIs.
- react-router-dom — for implementing seamless navigation between multiple pages or views.

5. Development Tools

Efficient development and collaboration are supported by the following tools and platforms:

- Code Editor: Visual Studio Code (VS Code), offering extensions for Python, React, and Git integration.
- Version Control: Git, used for maintaining code versions and enabling team collaboration via platforms like GitHub.
- API Testing Tool: Postman, for designing, testing, and debugging API endpoints before integration.

6. Machine Learning Model Deployment

The trained Linear Regression model is serialized for deployment within the Django backend:

- The model is saved using Pickle or Joblib for persistence.
- During prediction, the model is loaded dynamically in Django views to process real-time user input.
- For large-scale or production-level deployments, the model can be containerized and served independently using FastAPI, Flask, or AWS Lambda to enhance scalability and response time.

7. Troubleshooting and Support Resources

To maintain system reliability and ensure smooth debugging, the following resources and communities are referenced regularly:

Official Documentation:

- Django: <https://docs.djangoproject.com/>
- React: <https://reactjs.org/docs/>
- MySQL: <https://dev.mysql.com/doc/>
- Node.js: <https://nodejs.org/en/docs>

Developer Communities:

- Stack Overflow — for troubleshooting technical errors and code queries.
- GitHub Discussions — for library-specific issues and open-source collaboration.
- Developer forums and subreddits for additional support.

Testing Tools:

- Postman for backend API verification.
- Browser DevTools for inspecting and debugging frontend performance issues.

VII. FUTURE SCOPE

While the current version of the College Prediction System demonstrates strong performance and practical usability, there remains considerable scope for expanding its capabilities and enhancing its overall impact. Future developments



can focus on several dimensions — including the inclusion of new data domains, advanced analytics, and greater user accessibility — to make the system more intelligent, inclusive, and data-driven.

1. Expanding the Scope and Coverage

One major direction for future work is to broaden the system's applicability beyond the MHT-CET examination. The model can be extended to include other prominent entrance exams such as NEET (for medical admissions), JEE (for engineering streams), and even postgraduate admission processes. This would transform the platform into a multi-domain academic guidance tool, capable of serving students from diverse academic backgrounds and specializations. Such integration would make the system a one-stop platform for career counseling and college recommendation across multiple disciplines in higher education.

2. Model Optimization and Feature Enrichment

To enhance prediction precision, the machine learning framework can be upgraded by experimenting with more sophisticated algorithms such as Random Forests, Gradient Boosting, or Neural Networks. These methods can better capture complex, nonlinear relationships between variables compared to basic regression models.

Additionally, the inclusion of richer feature sets — such as gender, detailed reservation subcategories, high school performance, and socio-economic indicators — could make the predictions more personalized. By capturing these finer nuances, the system can deliver a more accurate and holistic evaluation of admission likelihood for each student profile.

3. Advanced Analytical Insights

A key enhancement could involve integrating probabilistic and interpretive analytics to provide students with more actionable insights. Instead of presenting a single predicted college, the system could classify outcomes into intuitive categories such as “Safe,” “Moderate,” and “Ambitious” based on confidence scores or probability distributions.

Furthermore, incorporating comparative analytics — allowing side-by-side comparison of multiple colleges based on factors like cutoff trends, placement records, and location preferences — would empower students to make informed, data-driven decisions rather than relying solely on rank-based suggestions.

4. Enhanced User Accessibility and Experience

Another significant area for improvement is user interface accessibility. To ensure inclusivity for students across different regions of India, the system could support multilingual interfaces (e.g., Marathi, Hindi, and other regional languages).

A particularly innovative addition could be the integration of optical character recognition (OCR) to automatically extract examination scores directly from uploaded mark sheets or PDFs. This would minimize manual data entry, reduce errors, and simplify the user experience, especially for first-time users.

5. Real-Time Data Integration and Dynamic Updates

The system's predictive accuracy could be further strengthened by integrating live data feeds from official government or university admission portals. Connecting to real-time seat matrix updates, counseling schedules, and round-wise cutoff trends would make the predictions dynamic and reflect the most current admission landscape. This would significantly enhance the model's reliability and relevance, particularly during active admission cycles when cutoff scores fluctuate rapidly.

VII. CONCLUSION

The comparative analysis of four key research studies on college admission prediction models reveals a diverse landscape of technological approaches — ranging from traditional data filtering and statistical regression to ensemble-based machine learning and deep neural network architectures. Each study demonstrates the growing importance of data-driven methods in improving admission guidance and decision support for students. A recurring finding across these works is that the inclusion of multidimensional features — such as academic scores, demographic details, reservation categories, and institutional attributes — significantly boosts predictive performance and reliability. These studies collectively highlight the transition from static, manually curated datasets to intelligent systems capable of learning complex admission patterns.

Despite these notable advancements, a closer evaluation indicates that many of the existing systems suffer from certain limitations in real-world applicability. For instance, while they may perform well in experimental environments, they



often fall short in areas like model interpretability, regional adaptability, and interactive usability. Most models provide only raw numerical predictions without clear explanations or dynamic contextual filtering (such as district-level or category-specific variations). Furthermore, the lack of real-time data integration limits their relevance during active admission cycles when cutoffs and seat matrices undergo frequent changes. Another significant shortcoming is the absence of user-centric design principles, making many systems difficult to navigate for students unfamiliar with technical interfaces or complex analytical outputs.

To address these challenges, our proposed project introduces an enhanced Dynamic College Prediction System that integrates both predictive intelligence and user experience design. Unlike previous approaches, this system enables district-wise, branch-wise, and category-wise filtering, offering more localized and personalized predictions aligned with Maharashtra's MHT-CET admission framework. The inclusion of intuitive PDF export functionality allows students to conveniently save and review personalized prediction reports offline, supporting informed decision-making even in areas with limited internet access.

By combining machine learning accuracy with practical usability, the proposed model not only fills the gap between prediction and interpretation but also democratizes access to admission insights. It empowers students—particularly those from rural or underrepresented regions—to make data-informed choices about their future education pathways with clarity, confidence, and convenience.

REFERENCES

- [1] A. Majeed Inamdar, T. Mhatre, P. Nadar and S. Joshi, "Personalized College Recommender and Cutoff Predictor for Direct Second Year Engineering," 2022 IEEE 7th International conference for Convergence in Technology (I2CT), Mumbai, India, 2022, pp. 1-4, doi: 10.1109/I2CT54291.2022.9825378.
- [2] P. N. Parkhi, A. Patel, D. Solanki, H. Ganwani and M. Anandani, "Machine Learning Based Prediction Model for College Admission," 2023 11th International Conference on Emerging Trends in Engineering & Technology - Signal and Information Processing (ICETET - SIP), Nagpur, India, 2023, pp. 1-6, doi: 10.1109/ICETET-SIP58143.2023.10151595.
- [3] Y. Zhang et al., "Research on the Prediction Method of the College Professional Admission Scores," 2022 International Seminar on Computer Science and Engineering Technology (SCSET), Indianapolis, IN, USA, 2022, pp. 406-409, doi: 10.1109/SCSET55041.2022.00098.
- [3] Y. Zhang et al., "Research on the Prediction Method of the College Professional Admission Scores," 2022 International Seminar on Computer Science and Engineering Technology (SCSET), Indianapolis, IN, USA, 2022, pp. 406-409, doi: 10.1109/SCSET55041.2022.00098.
- [4] Zhenru Wang and Yijie Shi, "Prediction of the admission lines of college entrance examination based on machine learning," 2016 2nd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, 2016, pp. 332-335, doi: 10.1109/CompComm.2016.7924718
- [5] K. Kumari, M. Kataria, V. Limbani, and R. Soni, "CAPSLG: College Admission Predictor and Smart List Generator," in Proc. 2nd Int. Conf. Advances Sci. Technol. (ICAST), Mumbai, India, Apr. 2019.
- [6] A. Pawar, R. Patil, K. Mathapati, P. Lonare, and L. Malphedwar, "College Admission Prediction Using Machine Learning," J. Emerg. Technol. Innov. Res., vol. 10, no. 11, Nov. 2023.
- [7] G. Raftopoulos, G. Davrazos, and S. Kotsiantis, "Fair and Transparent Student Admission Prediction Using Machine Learning Models," Algorithms, vol. 17, no. 12, Art. 572, Dec. 2024.
- [8] H. Lee, R. F. Kizilcec, and T. Joachims, "Evaluating a Learned Admission-Prediction Model as a Replacement for Standardized Tests in College Admissions," arXiv preprint arXiv:2302.03610, Feb. 2023.
- [9] A. Priyadarshini, B. Martinez-Neda, and S. Gago-Masague, "Admission Prediction in Undergraduate Applications: An Interpretable Deep Learning Approach," arXiv preprint arXiv:2401.11698, Jan. 20
- [10] M. M. Khalili, X. Zhang, M. Abroshan, and S. Sojoudi, "Improving Fairness and Privacy in Selection Problems," arXiv preprint arXiv:2012.03812, Dec. 2020.



- 11] A. J. Kehinde, A. E. Adeniyi, R. O. Ogundokun, H. Gupta, and S. Misra, "Prediction of Students Performance with Artificial Neural Network using Demographic Traits," arXiv preprint arXiv:2108.07717, Aug. 2021.
- 12] H. Lee, R. F. Kizilcec, and T. Joachims, "Evaluating a Learned Admission-Prediction Model as a Replacement for Standardized Tests in College Admissions," arXiv preprint arXiv:2302.03610, Feb. 2023.
- 13] M. S. A. Basha, C. Prabhavathi, V. Khangembam, M. M. Sucharitha, and P. M. Oveis, "Predicting Graduate Admissions using Ensemble Machine Learning Techniques: A Comparative Study of Classifiers and Regressors," in Proc. 2nd Int. Conf. Innovation in Technology (INOCON), 2023.
- 14] C. A. D. A., M. C. N., P. R., B. S. S., and S. S., "Prediction for University Admission using Machine Learning," Int. J. Recent Technol. Eng. (IJRTE), vol. 8, no. 6, Mar. 2020, doi: 10.35940/ijrte.F9043.038620.
- 15] S. Patel, H. Waghela, P. Gupta, and N. Rajgor, "Prediction of Graduate Admission Using Machine Learning," Int. J. Scientific Research in Comp. Sci. Engg. Info. Tech. (IJSRCSEIT), vol. 8, no. 5, Sep-Oct 2022, doi: 10.32628/CSEIT228534

