Journal Homepage: www.ijarpr.com ISSN: 3049-0103 (Online)

International Journal of Advance Research Publication and Reviews

Vol 03, Issue 11, pp 49-61, November, 2026

CET College Recommendation and Smart Option Form Generator using AI: A Review of Machine Learning, Historical Cutoff Analysis, and CAP Option Form Optimization

Harshal Ahire*, Pranav Magar*, Kadam Kunal*, Shriram Shinde*, Farendrakumar Ghodichor**

*Student 's, Dept. of AI and Data Science Dr. D. Y. Patil College of Engineering and Innovation, Varale, Pune, India

**Assistant Professor, Dept. of AI and Data Science, Dr. D. Y. Patil College of Engineering and Innovation, Varale, Pune, India

harshalahire7499@gmail.com, pranavmagar8102@gmail.com, kadam10kunal@gmail.com, shriramshinde0714@gmail.com,

farendrakumar.ghodichor@dypatilef.com

ABSTRACT

The increasing number of engineering colleges, academic specializations, and admission choices has made college selection a complex decision-making process for students appearing for entrance examinations. In Maharashtra, candidates participating in the Maharashtra Common Entrance Test (MHT-CET) Centralized Admission Process (CAP) must evaluate several factors, including percentile, category, applicable seat type, home university, branch preferences, institute location, fees, and historical admission cutoffs. Manual analysis of these factors can be time-consuming and may result in unsuitable or poorly ordered college preferences.

This review examines research areas relevant to artificial intelligence, machine learning, educational recommender systems, admission prediction, historical cutoff analysis, multi-criteria decision-making, and explainable decision-support systems. It organizes the field into recommendation, prediction, ranking, cutoff analysis, and explanation techniques, and outlines criteria for comparing their data requirements, interpretability, limitations, and applicability to MHT-CET admissions. The review methodology specifies literature databases, search queries, inclusion and exclusion criteria, and a data-extraction framework to support a reproducible survey.

Based on these themes, the paper presents candidate research gaps and a conceptual framework integrating student-profile analysis, verified historical CAP data, personalized college recommendations, indicative Dream–Target–Safe groupings, and automated CAP option-form ranking. The framework also considers changing admission rules, data provenance, privacy, fairness, uncertainty, and the need for temporal validation. The framework is a research direction rather than a validated system; its effectiveness must be established through a verified literature survey, implementation, baseline comparisons, and evaluation using reliable admission data. Historical compatibility and model predictions cannot guarantee future admission.

Keywords: Artificial Intelligence; Machine Learning; College Recommendation System; MHT-CET; Centralized Admission Process; Historical Cutoff Analysis; Educational Recommender Systems; Decision Support System; Multi-Criteria Decision-Making; Option Form Optimization; Explainable AI.

1. Introduction

1.1 Background

The growth of higher education has increased the number of engineering institutions, academic programs, and specializations available to students. Selecting an appropriate engineering college is an important decision that can influence academic development, career opportunities, financial commitments, and future professional goals. Making this decision requires evaluating several parameters simultaneously, which can make the admission process challenging.

In India, engineering admissions are conducted through entrance examinations and centralized counselling procedures. In Maharashtra, the Maharashtra Common Entrance Test (MHT-CET) is an important entrance examination for admission to undergraduate engineering and technology programs. Eligible

candidates participate in the applicable Centralized Admission Process (CAP), where they prepare and submit preferences for institutes and courses according to the rules for the relevant admission year.

Students may need to consider examination percentile, category, applicable seat type, home university, preferred engineering branch, institute location, fees, and previous admission cutoffs. Understanding the relationship between these factors requires effort, particularly when comparing many colleges and branches across multiple CAP rounds.

1.2 Problem Statement

Traditional college-selection approaches often depend on manually searching college websites, examining previous-year cutoff lists, consulting seniors, and obtaining guidance from teachers or admission counsellors. These methods can provide useful information, but they may not offer a unified and personalized analysis of all relevant factors.

Historical cutoff information can help students understand previous admission trends, but comparisons require attention to admission year, category, seat type, institute, branch, and CAP round. A cutoff achieved in a previous year does not guarantee the same outcome in a future admission cycle. There is therefore a need to investigate decision-support approaches that organize admission data, compare college options, provide personalized recommendations, and assist students in preparing a well-informed CAP option form.

1.3 Role of Artificial Intelligence and Machine Learning

Artificial intelligence (AI) and machine learning (ML) provide techniques for analyzing structured datasets, identifying patterns, estimating outcomes, and supporting data-driven decisions. These techniques have potential applications in educational recommendation systems, student guidance, admission analysis, and personalized information retrieval.

Rule-based filtering removes options that do not satisfy specified eligibility or preference constraints. Content-based recommendation suggests colleges based on similarities between student preferences and college characteristics. Collaborative filtering can use patterns in user–item interactions when suitable interaction data exists. Machine-learning models may estimate a defined admission-related outcome when representative labelled data is available. Multi-criteria ranking combines factors such as branch preference, location, fees, and historical cutoff compatibility. Explainable AI can help communicate why a particular college or branch has been recommended.

These methods are complementary rather than interchangeable. A ranking algorithm can organize preferences without predicting admission outcomes, while a predictive model requires suitable training data and validation before its estimates can be trusted.

1.4 Motivation

The motivation for this review is to examine how research in recommendation systems, educational data mining, machine learning, and decision-support systems can inform college selection. The Maharashtra CAP context is particularly relevant because students must organize multiple institute and course preferences under specific admission rules. A useful decision-support framework could combine historical cutoff analysis with student preferences and transparent ranking criteria, helping students compare options systematically.

The review also considers data reliability, changing admission rules, model interpretability, privacy, and the limitations of predictions based on historical information. These considerations are essential when designing systems that may influence students' educational decisions.

1.5 Objectives

The objectives of this review are:

1. To examine research on college recommendation systems and educational recommender systems.

2. To review machine-learning techniques used for prediction, classification, ranking, and decision support.

3. To study how historical admission cutoffs can support college-selection decisions.

4. To compare existing approaches based on data requirements, evaluation, interpretability, scalability, and limitations.

5. To identify research gaps relevant to personalized MHT-CET recommendations and CAP option-form preparation.

6. To outline a potential framework integrating student preferences, historical CAP data, recommendation techniques, and transparent option ranking.

7. To discuss challenges related to data quality, admission-rule changes, fairness, privacy, and real-world validation.

1.6 Scope and Intended Contributions

This paper focuses on reviewing existing methods rather than claiming the successful implementation of a new admission prediction model. It covers recommendation techniques, historical data analysis, machine-learning approaches, ranking strategies, and decision-support mechanisms relevant to engineering college admissions.

The intended contributions are a structured classification of relevant techniques; a comparative framework for their applications and limitations; candidate research gaps to verify through the literature survey; a conceptual framework for MHT-CET college recommendation and CAP option-form assistance; and a discussion of evaluation requirements for a reliable, transparent, student-oriented system. The framework remains conceptual until implemented and empirically evaluated.

2. Background and Problem Definition

2.1 MHT-CET and the Centralized Admission Process

The Maharashtra Common Entrance Test (MHT-CET) is used for admission to various undergraduate professional courses in Maharashtra. For engineering admissions, eligible candidates participate in the applicable CAP, subject to the eligibility requirements and rules for that admission year. Candidates consider examination performance, category, seat type, home university, branch preference, location, fees, and previous cutoffs.

The process involves identifying suitable institutes, comparing courses, preparing an ordered preference list, and responding to later rounds when applicable. The specific process, deadlines, seat types, and rules must be checked against the official information brochure for the relevant admission year.

2.2 Historical Cutoff Analysis

Historical cutoff data records admission-related thresholds for particular institutes, courses, admission categories, seat types, and rounds. When interpreted correctly, it can help students understand previous admission trends. However, cutoffs can change year to year; categories and seat types may have different cutoffs; CAP rounds may produce different outcomes; and changes in seat availability, applicant preferences, and admission rules can affect results. Historical cutoffs should therefore be treated as decision-support evidence, not as a guarantee of admission.

2.3 College Recommendation Systems

A college recommendation system helps users identify and compare institutions according to their profiles and preferences. Relevant approaches include:

• Rule-based filtering: applies explicit eligibility and preference rules.

• Content-based recommendation: compares student preferences with college and course attributes.

• Collaborative filtering: uses patterns in student–college interactions when suitable data exists.

• Weighted scoring: combines several criteria using defined weights.

• Machine-learning prediction: estimates outcomes from suitable historical examples.

• Hybrid recommendation: combines methods to address individual limitations.

Rule-based filtering and weighted ranking provide useful baselines for this application. More complex predictive models should be included only when reliable data and appropriate evaluation are available.

2.4 Formal Problem Definition

The central problem is the difficulty students face when converting examination results, admission rules, historical cutoff information, and personal preferences into an informed and logically ordered college preference list. Relevant information may be distributed across official documents, institute websites, cutoff tables, and counselling resources. Students may struggle to compare alternatives consistently, understand the risks of relying on historical cutoffs, and update preferences as new information becomes available.

This review investigates how recommendation systems, historical data analysis, machine learning, and multi-criteria decision-making can be combined to support college selection and CAP option-form preparation.

3. Review Methodology

3.1 Review Design

The paper proposes a structured literature-review methodology. The objective is to identify, screen, classify, and compare research relevant to college recommendation, educational recommender systems, admission prediction, and decision-support methods. The work should not be described as a completed systematic review unless reproducible searches, screening, and a complete study-selection record are maintained.

3.2 Literature Sources

Potential sources include Google Scholar, IEEE Xplore, SpringerLink, ScienceDirect, and the ACM Digital Library. Official Maharashtra CET Cell information brochures and CAP documents should be used to verify admission procedures and terminology rather than relying exclusively on secondary websites.

3.3 Search Queries

Suggested queries include:

• “college recommendation system” AND “machine learning”

• “educational recommender system” AND “higher education”

• “university admission prediction” AND “machine learning”

• “student decision support system” AND “college selection”

• “explainable AI” AND “education”

• “multi-criteria decision making” AND “university selection”

• “engineering admission” AND “cutoff prediction”

• “MHT-CET” AND “CAP admission”

During the actual search, record the search date, exact query, filters, database, and result count.

3.4 Inclusion and Exclusion Criteria

Include peer-reviewed research on recommendation and decision-support methods, relevant machine-learning prediction studies, educational data-mining research, papers that explain evaluation methods and limitations, and official admission documents relevant to MHT-CET CAP. Exclude unverifiable publications, fabricated references, unrelated studies, duplicate records, and sources that lack sufficient methodological information. Outdated admission documents should not be used to describe current rules without qualification.

3.5 Data Extraction and Comparative Analysis

For each selected study, record authors, publication year, title, research objective, dataset and source, method, evaluation metrics and validation approach, reported findings in their original context, author-identified limitations, and relevance to MHT-CET college recommendation. Group the studies by theme and compare them critically. Performance values should not be compared directly when datasets, target variables, validation procedures, or evaluation metrics differ substantially.

4. Literature Survey: Thematic Review

4.1 Educational Recommender Systems

Recommender systems help users identify relevant options from large collections of items. In education, they can support course selection, learning-resource discovery, academic planning, and institution selection. Traditional approaches include content-based filtering, collaborative filtering, and hybrid systems. Content-based methods use item characteristics and user preferences; collaborative filtering uses patterns in interactions between users and items; and hybrid methods combine strategies to address limitations associated with individual methods.

For college selection, content-based recommendation can compare student preferences with attributes such as engineering branch, location, fees, and institute type. Collaborative filtering requires appropriate student–college interaction data and careful consideration of privacy and representativeness.

4.2 Machine Learning for Admission-Related Prediction

Machine-learning models learn relationships between input features and a defined outcome from historical examples. Depending on available data, the target might be an admission outcome, a cutoff category, or another measurable variable. Potential algorithms include logistic regression, decision trees, random forests, and gradient-boosting methods. Algorithm selection should depend on the prediction task, data quality, interpretability requirements, and validation results.

A model trained on historical outcomes may not generalize well when admission rules, seat availability, or student preferences change. Temporal validation—training on earlier admission years and testing on a later year—is particularly relevant.

4.3 Historical Cutoff Analysis

Cutoff analysis examines how admission thresholds vary across institutions, branches, categories, seat types, years, and rounds. Descriptive statistics and visualization can reveal historical patterns, while predictive techniques may estimate future values when sufficient, consistent data is available. A cutoff comparison tool can work without ML: it can calculate the difference between a student's percentile and a comparable historical cutoff, provided both values use the same scale and relevant admission conditions.

This difference must not be presented as a calibrated probability of admission. A model estimating admission probability requires suitable outcome labels, representative data, calibration, and independent validation.

4.4 Multi-Criteria Decision-Making

College selection is a multi-criteria decision problem because students may prioritize different combinations of branch, affordability, location, institute characteristics, and historical admission compatibility. Weighted scoring can combine criteria into a transparent ranking:

$$S(c) = \Sigma_{j=1 \text{ to } m}\, w_j\, r_j(c)$$

Here, S(c) is the score of college-course option c, w_j is the weight assigned to criterion j, and $r_j(c)$ is its normalized score. Weights should be defined transparently and ideally informed by student preferences. Eligibility requirements should be treated as constraints rather than criteria that can be compensated for by a high score elsewhere.

4.5 Explainable AI and Decision Support

Explainability matters when recommendations influence educational choices. A student should understand why an option appears in a list and which factors affect its ranking. Explanations may state that a course matches the preferred branch, an institute satisfies a location preference, fees fall within the budget, or historical cutoff information is compatible with the student's profile under the selected comparison criteria.

These explanations can be generated from explicit rules and ranking components. Complex predictive models may require additional explanation methods, but explanations should remain faithful to the underlying model.

4.6 Literature Survey Extraction Table

The table below is a structure for recording verified studies, not a claim that these studies have already been screened. Populate it with actual papers before submission.

No.	Reference / year	Research focus	Main contribution / reported scope	Limitations / relevance to this review
1	Resnick & Varian (1997), “Recommender Systems”	Recommender systems overview	Introduces the recommender-systems problem and its role in helping users navigate large information spaces.	Foundational overview; not specific to college admissions.
2	Adomavicius & Tuzhilin (2005), “Toward the Next Generation of Recommender Systems”	Survey of recommender techniques	Surveys state-of-the-art recommendation approaches and discusses extensions beyond basic methods.	General survey; admission-specific data and rules require separate treatment.

No.	Reference / year	Research focus	Main contribution / reported scope	Limitations / relevance to this review
3	Burke (2002), “Hybrid Recommender Systems: Survey and Experiments”	Hybrid recommendation	Examines hybrid recommendation strategies that combine methods to address limitations of individual approaches.	Performance depends on component methods and evaluation context.
4	Koren, Bell & Volinsky (2009), “Matrix Factorization Techniques for Recommender Systems”	Collaborative filtering / latent factors	Explains matrix-factorization methods for learning latent user–item relationships in recommendation tasks.	Requires meaningful interaction data; not directly transferable to admission cutoffs.
5	Ricci, Rokach & Shapira (eds.) (2011/2015), Recommender Systems Handbook	Recommender systems reference	Provides broad coverage of recommendation algorithms, evaluation, and system design.	Book/handbook rather than a single empirical college-admission study; edition must be specified.
6	Pedregosa et al. (2011), “Scikit-learn: Machine Learning in Python”	Machine-learning tools	Describes a Python machine-learning library supporting model selection, preprocessing, and evaluation workflows.	Software/library paper; does not establish admission prediction performance.
7	Chen & Guestrin (2016), “XGBoost: A Scalable Tree Boosting System”	Gradient boosting	Presents a scalable tree-boosting system and its design for machine-learning tasks.	General algorithm paper; suitability for CAP data must be tested empirically.
8	Ribeiro, Singh & Guestrin (2016), “Why Should I Trust You?: Explaining the Predictions of Any Classifier”	Model explanation (LIME)	Introduces local explanations for individual classifier predictions.	Explanations approximate model behavior and should not be treated as proof of causal reasoning.
9	Lundberg & Lee (2017), “A Unified Approach to Interpreting Model Predictions”	Model explanation (SHAP)	Presents a unified feature-attribution framework for interpreting model predictions.	Feature attribution does not itself guarantee fairness or model correctness.
10	Breiman (2001), “Random Forests”	Ensemble classification/regression	Introduces random forests, an ensemble method based on randomized decision trees.	General methodology; requires task-specific validation and careful interpretation.
11	Friedman (2001), “Greedy Function Approximation: A Gradient Boosting Machine”	Gradient boosting	Develops gradient boosting as a function-approximation approach using additive models.	General ML method; model complexity and overfitting need control.
12	Schafer, Konstan & Riedl (2001), “E-Commerce Recommendation Applications”	Recommendation system applications	Discusses recommender-system applications and design issues in e-commerce contexts.	Different application domain; use for general design concepts only.

5. Comparative Analysis

5.1 Comparison of Recommendation and Prediction Techniques

The following comparison summarizes general methodological characteristics. It is not an experimental ranking of algorithms.

Technique	Main purpose	Advantages	Limitations
Rule-based filtering	Apply eligibility and preference constraints	Transparent; easy to audit	Requires accurate, updated rules
Content-based recommendation	Match preferences with college attributes	Does not require a large interaction dataset	Depends on item features and preference quality
Collaborative filtering	Use user–item interaction patterns	May identify latent preference patterns	Needs interaction data; cold-start and sparsity
Weighted scoring	Rank using multiple criteria	Transparent and customizable	Depends on weights and normalization
Logistic regression	Estimate a defined outcome probability	Interpretable baseline	May miss complex relationships
Random forest	Model nonlinear patterns	Flexible for structured data	Less transparent than simple rules
Gradient boosting	Model complex structured-data patterns	Potentially strong performance on suitable data	Needs tuning and overfitting control
Explainability methods	Explain model outputs	Can support understanding and auditing	Quality depends on method and model

Module	Suitable evaluation measures
Admission classification	Precision, recall, F1-score, confusion matrix
Probability prediction	Brier score, log loss, calibration analysis
Cutoff forecasting	MAE, RMSE, comparison with baseline forecasts
Ranking	Precision@K, Recall@K, NDCG@K when ground truth exists
Data processing	Missing-value rate, duplicate rate, consistency checks
User experience	Task completion, usability feedback, user-study results

5.2 Evaluation Metrics

Different modules require different metrics. Admission classification can be assessed with precision, recall, F1-score, and a confusion matrix. Probability prediction can use Brier score, log loss, and calibration analysis. Numerical cutoff forecasting can use mean absolute error (MAE), root mean squared error (RMSE), and baseline comparisons. Ranking can use Precision@K, Recall@K, and normalized discounted cumulative gain (NDCG@K) when

appropriate ground truth exists. Data processing can be evaluated using missing-value rate, duplicate rate, and consistency checks. User experience can be assessed using task completion, usability feedback, and user-study results.

Accuracy alone is insufficient, particularly when classes are imbalanced. A ranking system cannot be validated merely by reporting classification accuracy. A fair evaluation should compare against simple baselines, use held-out data, and document the dataset, preprocessing, features, and evaluation procedure.

5.3 Comparison Principles

Compare approaches on: (1) data requirements and availability; (2) interpretability; (3) personalization; (4) temporal reliability on later admission years; (5) fairness and representation across relevant categories and seat types; (6) practical usefulness and decision quality; and (7) maintainability as data and rules change. No method should be described as more accurate or superior without evidence from comparable experiments.

6. Research Gaps

6.1 Integration of Multiple Admission Factors

A candidate gap is the limited integration of student preferences, historical cutoff compatibility, affordability, location, and admission constraints in one coherent decision-support workflow. The literature review must establish whether existing systems already integrate these elements and what limitations remain.

6.2 Round-Wise Admission Support

Admission choices may change as new cutoff or vacancy information becomes available. A useful research direction is to examine systems that update recommendations using new, verified data while preserving the relevant rules for each round.

6.3 Transparent Risk Grouping

Dream, Target, and Safe labels can help students organize options, but their criteria must be clearly defined. A research opportunity is to compare transparent rule-based grouping with validated probabilistic methods while making clear that neither guarantees admission.

6.4 Automated Option-Form Ranking

Recommendation and option-form generation are related but distinct tasks. A recommendation system identifies relevant choices; an option-form assistant must also respect official procedures, eligibility requirements, user priorities, and ordering. A possible contribution is to translate recommendation scores into a user-editable preference list with explanations and validation checks.

6.5 Explainability and User Trust

A technically reasonable ranking may still be difficult to understand. Research could investigate whether explanations improve user understanding and decision quality.

6.6 Data Quality and Temporal Validation

Historical data may contain inconsistent institute names, missing values, different cutoff representations, and changes in admission rules. A robust framework should address these issues and evaluate performance on later admission years.

6.7 Fairness, Privacy, and Responsible Use

Admission-related data may contain sensitive personal information. The system should minimize data collection, control access, document the use of category-related attributes, and avoid presenting uncertain outputs as authoritative decisions. These are candidate research gaps to verify against the literature, not established findings of a completed systematic review.

7. Proposed Framework and Architecture

7.1 High-Level Workflow

The conceptual workflow consists of: (1) student information and preferences; (2) historical CAP data processing and validation; (3) recommendation and ranking methods; (4) explanation and option-form assistance; and (5) student review and export. The diagram below summarizes the proposed workflow.

Stage	Inputs / processing	Outputs
1. Student interface	Percentile, category, branch, location, budget, preferences	Structured student profile
2. Data acquisition and validation	Official CAP documents, cutoff data, institute details	Clean, traceable data
3. Recommendation engine	Rules, historical comparison, scoring, optional validated ML	Candidate college-course options
4. Explanation and option-form assistance	Ranking rationale, indicative groupings, validation	Explained and ordered recommendations
5. Student review and export	User edits, reorder, final checks	User-reviewed suggested list

7.2 Data Processing Layer

The data-processing layer should collect and standardize information from reliable sources. Each cutoff record should preserve admission year and CAP round, institute identifier and name, course or branch, category and applicable seat type, cutoff value and representation, and data source and retrieval date. The system should check duplicate entries, missing values, inconsistent names, and incompatible cutoff formats. Records should not be combined merely because institute or branch names appear similar. Data provenance should allow administrators to trace recommendations to underlying source records.

7.3 Recommendation Engine

Stage 1—Constraint filtering: apply applicable eligibility and admission rules and distinguish mandatory rules from optional preferences.

Stage 2—Historical comparison: compare the student's profile with comparable historical cutoff records, displaying the year, round, category, and seat type used.

Stage 3—Preference ranking: rank eligible or potentially relevant options using configurable weights for branch, location, affordability, and other preferences.

Stage 4—Optional prediction: if a suitable labelled dataset exists, evaluate an ML model for a clearly defined outcome. Communicate uncertainty and validation limitations rather than guaranteeing admission.

7.4 Dream–Target–Safe Classification

These labels should be indicative planning groups, not official admission categories. Dream may indicate a more ambitious option relative to the selected historical comparison; Target may indicate comparatively compatible historical admission data; and Safe may indicate more favourable historical compatibility under the chosen criteria. Thresholds must be defined and justified. Safe must never mean guaranteed admission. Grouping should account for differences in rounds, seat types, and categories.

7.5 Suggested Technical Architecture

A possible implementation uses React and Vite for the frontend; Tailwind CSS for responsive styling; Node.js and Express.js for APIs, authentication, and application logic; MongoDB for profiles, institute data, and preferences; Python with Pandas and NumPy for data cleaning and analysis; Scikit-learn with Flask or FastAPI for an optional ML service; Joblib or an appropriate model format for versioned model persistence; Chart.js for visualizations; and an optional language-model API for counselling.

These are implementation choices, not findings established by the review. The engine can initially use deterministic rules and scoring and be extended if validated ML models provide measurable benefits.

Component	Possible technology	Responsibility
Frontend	React + Vite	Profile entry, dashboard, college comparison
Styling	Tailwind CSS	Responsive UI
Backend	Node.js + Express.js	APIs, authentication, application logic
Database	MongoDB	Profiles, institute data, saved preferences
Data analysis	Python, Pandas, NumPy	Cleaning and historical analysis
ML service	Scikit-learn + Flask/FastAPI	Optional model training and inference
Model storage	Joblib or suitable format	Versioned model persistence
Charts	Chart.js or similar	Cutoff trends and comparisons
AI counselling	Optional language-model API	Question answering and explanations

7.6 Conceptual Data Model

Suggested entities are: StudentProfile (relevant admission attributes and preferences); Institute (identifier, name, location and verified details); Course (course or branch associated with an institute); CutoffRecord (year, round, category, seat type, cutoff and source); Recommendation (selected criteria, ranking score, explanation and data version); OptionList (ordered institute-course preferences and user modifications); and ModelVersion (training-data version, features, validation results and deployment information). Sensitive information should be collected only when necessary, with appropriate access controls, retention policies, and secure handling.

Entity	Key information
StudentProfile	Relevant admission attributes and preferences
Institute	Identifier, name, location, verified details
Course	Course/branch associated with institute
CutoffRecord	Year, round, category, seat type, cutoff, source
Recommendation	Criteria, score, explanation, data version
OptionList	Ordered institute-course preferences and user edits
ModelVersion	Training-data version, features, validation and deployment details

8. Challenges and Future Scope

8.1 Data Availability and Quality

Admission data may be distributed across documents and may contain inconsistencies in naming, formatting, or cutoff representation. Automated extraction can introduce errors, so records should be validated against original sources. Future work should develop a reproducible data pipeline with

source tracking, validation rules, duplicate detection, and controlled dataset updates.

8.2 Changing Admission Rules

Admission policies, seat availability, reservation rules, and institutional offerings can change between years. A model trained on older data may become less reliable. Future work should maintain versioned admission rules, evaluate models on later-year data, and warn users when current conditions differ from training data.

8.3 Prediction Uncertainty

A historical cutoff is not a direct measure of an individual student's admission probability. Changes in applicant preferences, available seats, and admission rounds can influence future outcomes. Future work should compare historical compatibility methods with validated predictive models, measure calibration when probabilities are produced, and communicate uncertainty.

8.4 Fairness and Privacy

Category and other personal attributes may be relevant to applicable admission rules, but their use must be restricted to legitimate purposes and handled carefully. Student information must not be unnecessarily exposed or reused. Future work should conduct privacy reviews, document feature usage, test systematic errors, and explain how profile attributes affect recommendations.

8.5 Explainability

Complex models may make it difficult for students to understand why an option ranks high or low. Future work should provide explanations based on verified data, display criteria contributing to scores, and allow users to adjust preferences and inspect ranking changes.

8.6 Real-World Evaluation

A recommendation system can appear useful without improving decision quality. Future work should evaluate the system against simple baselines, assess ranking quality using appropriate ground truth, conduct usability studies, and compare results across admission years. Improvements should only be reported when measured.

8.7 AI Counselling

A language-model-based counselling assistant could explain system outputs and answer questions, but generated answers may be incorrect or outdated. Future work should ground answers in verified admission documents, cite relevant sources and years, distinguish official rules from general guidance, and direct users to official sources for consequential decisions.

9. Conclusion

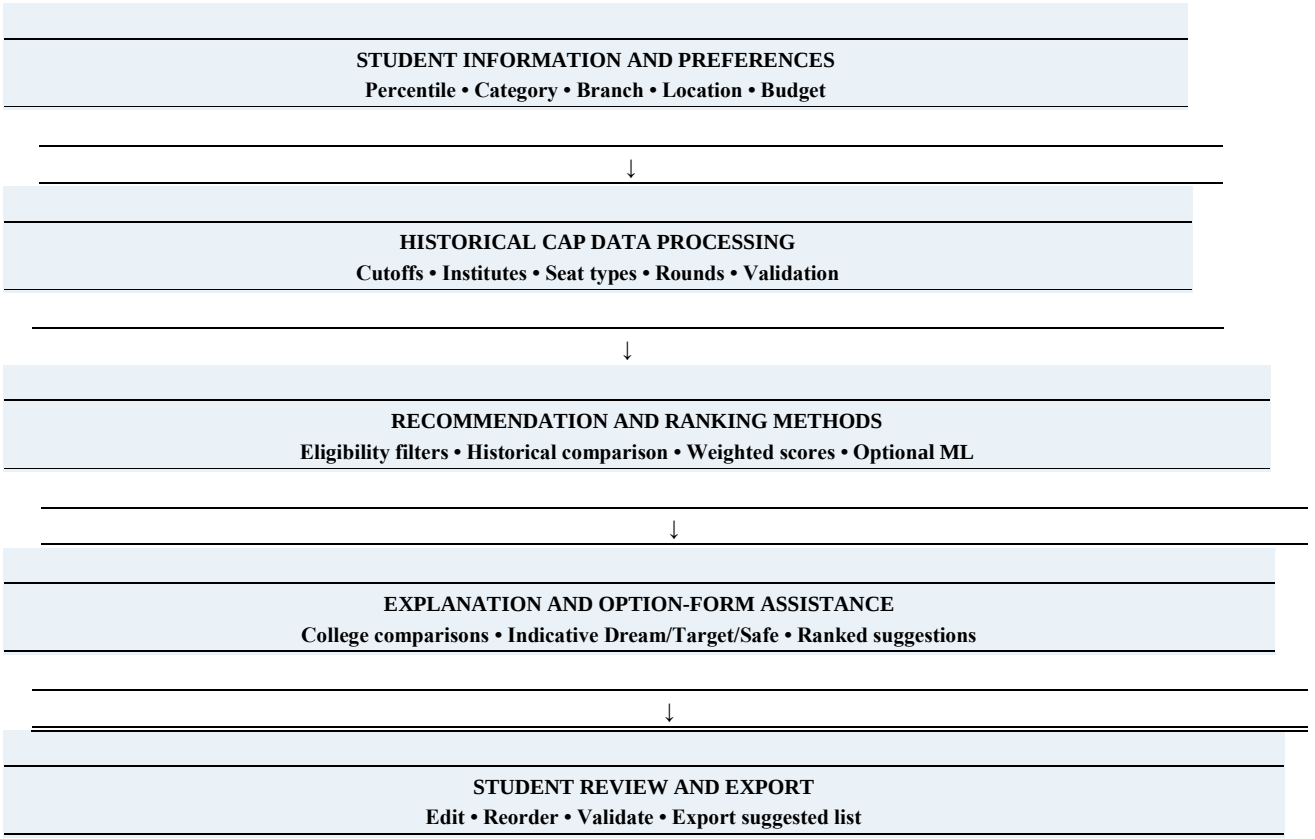
Selecting an appropriate engineering college requires students to evaluate examination performance, admission eligibility, preferred branch, institute location, affordability, and historical admission trends. In the MHT-CET CAP, interpreting cutoff information and preparing a suitable preference list can be challenging when information is distributed across multiple sources.

This review organizes relevant research areas involving educational recommender systems, machine learning, historical cutoff analysis, multi-criteria decision-making, and explainable AI. Rule-based filtering and transparent ranking can provide a practical foundation, while machine-learning techniques may add predictive capabilities when reliable labelled data and suitable validation are available.

A conceptual framework has been outlined to combine student preferences, verified admission data, historical comparisons, explainable recommendations, and CAP option-form assistance. The framework emphasizes data quality, transparency, privacy, uncertainty, and the distinction between historical compatibility and guaranteed admission.

The main research opportunity is to determine whether integrating these techniques can provide more understandable, consistent, and useful admission guidance than existing manual or single-method approaches. This must be investigated through a verified literature survey, comparison with established baselines, and evaluation using reliable admission data. Future work should complete the literature survey, verify research gaps, implement a reproducible prototype, and assess performance and usability. The proposed framework is a research direction rather than a validated solution; its effectiveness must be established empirically.

Figure 1. Conceptual Recommendation Workflow



10. References

1. P. Resnick and H. R. Varian, "Recommender Systems," Communications of the ACM, vol. 40, no. 3, pp. 56–58, 1997.
2. G. Adomavicius and A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 6, pp. 734–749, 2005.
3. R. Burke, "Hybrid Recommender Systems: Survey and Experiments," User Modeling and User-Adapted Interaction, vol. 12, pp. 331–370, 2002.
4. Y. Koren, R. Bell, and C. Volinsky, "Matrix Factorization Techniques for Recommender Systems," Computer, vol. 42, no. 8, pp. 30–37, 2009.
5. F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
6. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.
7. M. T. Ribeiro, S. Singh, and C. Guestrin, "Why Should I Trust You?: Explaining the Predictions of Any Classifier," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1135–1144.
8. S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in Advances in Neural Information Processing Systems, vol. 30, 2017.
9. Adomavicius, N. Manouselis, and Y. K. Kwon, "Multi-Criteria Recommender Systems," in Recommender Systems Handbook, 2nd ed., F. Ricci et al., Eds. Springer, 2015, pp. 769–803.
10. Ricci, L. Rokach, and B. Shapira, Eds., Recommender Systems Handbook, 2nd ed. New York, NY, USA: Springer, 2015.

11. Breiman, “Random Forests,” Machine Learning, vol. 45, pp. 5–32, 2001.

12. H. Friedman, “Greedy Function Approximation: A Gradient Boosting Machine,” Annals of Statistics, vol. 29, no. 5, pp. 1189–1232, 2001.

13. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, “GroupLens: An Open Architecture for Collaborative Filtering of Netnews,” in Proc. ACM CSCW, 1994, pp. 175–186.

14. Koren, “Collaborative Filtering with Temporal Dynamics,” in Proc. ACM SIGKDD, 2009, pp. 447–456.

15. Burke, “The Adaptive Web: Hybrid Web Recommender Systems,” in The Adaptive Web, Lecture Notes in Computer Science, vol. 4321. Springer, 2007, pp. 377–408.

16. 16.A. Konstan and J. Riedl, “Recommender Systems: From Algorithms to User Experience,” User Modeling and User-Adapted Interaction, vol. 22, pp. 101–123, 2012.

17. J. Delahoz-Domínguez and R. Hijón-Neira, "Recommender System for University Degree Selection: A Socioeconomic and Standardised Test Data Approach," Applied Sciences, 2024.

18. Mandavgane, S. Vaidya, H. Rahate, V. Khatake, and V. Naikwadi, "AI-Driven System for Engineering College Admission Prediction & Allocation," 2025.

19. Deschênes, "Recommender Systems to Support Learners’ Agency in a

20. Learning Context: A Systematic Review," International Journal of Educational Technology in Higher Education, 2020.

21. Algarni and F. Sheldon, "Systematic Review of Recommendation

22. Systems for Course Selection," Machine Learning and Knowledge Extraction, 2023.

23. Kumar, G. K. Kazoba, A. Kumar, H. Sharma, U. S. Yadav, and A.

24. Gupta, "An ML-Driven Decision Support System for Engineering College Selection Post-Intermediate Education in India," Discover Global Society, 2026.

25. R. Ambole, K. Hinge, P. Jatiya, R. Bhosale, and S. Bolshette, "College Recommendation System Using Machine Learning," 2023.

26. H. Motiramani, V. Melkari, and M. Mehta, "A Hybrid Recommendation System for University Selection Using Machine Learning," 2025.

27. M. Elahi, A. Starke, N. El Ioini, A. A. Lambrix, and C. Trattner, "Developing and Evaluating a University Recommender System," Frontiers in Artificial Intelligence, 2022.

28. C. Tenison, G. Ling, and L. McCulla, "Supporting College Choice Among International Students through Collaborative Filtering," International Journal of Artificial Intelligence in Education, 2023.

Additional sources required before submission: verified studies specifically addressing university/college recommendation; admission or student-outcome prediction; educational decision-support and explainable recommendation; and official Maharashtra CET Cell CAP documents for the relevant admission year. Verify bibliographic details against publisher records and add in-text citations. The references above support general technical background and do not constitute a complete MHT-CET-specific literature survey.