International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** X **Month of publication:** October 2025

DOI: https://doi.org/10.22214/ijraset.2025.74699

EduAdvisor-Academic and Career Planning using Decision Tree Algorithm

T.K.Dharshini[1], S.Dhivya[2], M.Nithya[3]
[1,2]Department of Computer Science Engineering, [3]Assistant Professor/CSE, KLN College of Engineering, Pottapalayam Sivagangai

***Abstract:** One ofthemostimportantstages forstudentsfinishing their12thgradeisacademicandcareerpreparation,buttraditional coaching approaches frequently lack scalability, personalization, and data-driven insights. Using optical character recognition (OCR)technologyandmachinelearningtechniques,EduAdvisorisacompletewebbasedprogramthatoffersindividualized,intelligentcareeradvice.Engineering,healthcare, artsandscience,business,science,andlawarethesixkeyjobcategoriesforwhichthesystemanalyzes academic achievement. It also incorporates decision tree algorithms, automatic marksheet processing, and secure user identification. The platform wasdeveloped withPython and theFlaskframework, andaccording to userinput, itachieves94.2%OCRaccuracy and 87.3% suggestion accuracy. The technology automatically analyzes academic papers, produces tailored suggestions with thorough justification, and offers thorough career exploration from the perspectives of the past, present, and future. EduAdvisor maintains excellent accuracy and user satisfaction scores of 4.5/5.0 while addressing the accessibility and scalability issues of traditional career advisingthroughadaptivelearning methodsandconstant feedback integration.ThisstudyshowshowAI-powerededucationaladvising systems can democratize access to high-quality materials for career preparation.*

***Keywords—**CareerGuidance,DecisionTreeAlgorithm,MachineLearning,OpticalCharacterRecognition,AcademicPlanning,Flask Framework, Educational Technology*

I. INTRODUCTION

One of a student's most important life decisions is the move from secondary school to college and a professional career [1], [2]. Although they are useful, traditional career counseling systems have serious problems with personalization, scalability, and accessibility [3], [4]. There is frequently a shortage of counselors in educational institutions, which leaves pupils without enough individualized care [5]. Geographical differences exacerbate these problems even more, as students in underserved or rural places have less access to resources that offer high-quality career counseling [6], [7]. Guidance systems must be able to handle large volumes of current data and deliver fast, pertinent recommendations due to the quick changes in career markets and the rise of new professional disciplines [8], [9]. Manual career counseling techniques, which rely on the availability and experience of individual counselors, are unable to effectively serve the expanding student body while upholding constant standards of quality [10]. Furthermore, systematic processes for monitoring results and continuously enhancing suggestion accuracy are frequently absentfrom old methodologies [11], [12].

Because machine learning technologies allow for automated, data-driven examination of academic performance patterns and career alignment,they present intriguing answers to these problems [13], [14]. Specifically, decision tree algorithms offer comprehensible outcomesthatassiststudentsincomprehendingtherationalebehindcareersuggestions,fosteringconfidenceandfaithintheirchoices [4]. By doing away with manual data entry and facilitating quick processing of academic documents, the incorporation of optical characterrecognitiontechnologysignificantlysimplifiestheprocedure[6],[15]. Throughanintegratedplatformthatblendsmachine learning-driven suggestions, OCR-based document processing, and thorough career information distribution, EduAdvisor tackles these issues [16], [17]. By offering individualized insights based on each user's unique academic background, the system seeks to democratize accesstohigh-quality career coaching[18].The design, implementation, andevaluationofEduAdvisor arepresentedin this study, showcasing how well it addresses the drawbacks of conventional career counseling while maintaining high accuracy and user happiness [19], [20].

II. SYSTEM ARCHITECTUREANDMETHODOLOGY

A. Overall System Architecture

Asillustrated in Fig. 1, EduAdvisor has a three-tier architecture that divides the presentation, businesslogic, anddata accesslevels [21], [22].

This modular design permits independent creation and testing of system components while guaranteeing security, scalability,and maintainability [23].Using HTML5,CSS3, and JavaScript,the presentation layer createsa responsive web interface that worksconsistentlyondesktop andmobiledevices[24].Through interactivevisualizationsandintuitivenavigation, the interface presents complex academic and career information with a focus on usability [25].User authentication, OCR processing, decision tree analysis, and suggestion creation are among the essential system functions that are coordinated by the business logic layer, which is based on the Flask framework [11]. This layer carries out data validation, security controls, and specialized service integration. Machine learning inference and OCR processing are two examples of compute-intensive tasks that may be scaled independently thanks to microservice design [26].

The data access layer stores user profiles, academic records, and career information in an ACID-compliant manner using PostgreSQLforproductiondeploymentandSQLitefordevelopment[27],[28].Referentialconstraintsandvalidationrulespreserve data integrity while supporting the sophisticated queries needed for recommendation creation in the database design [29].

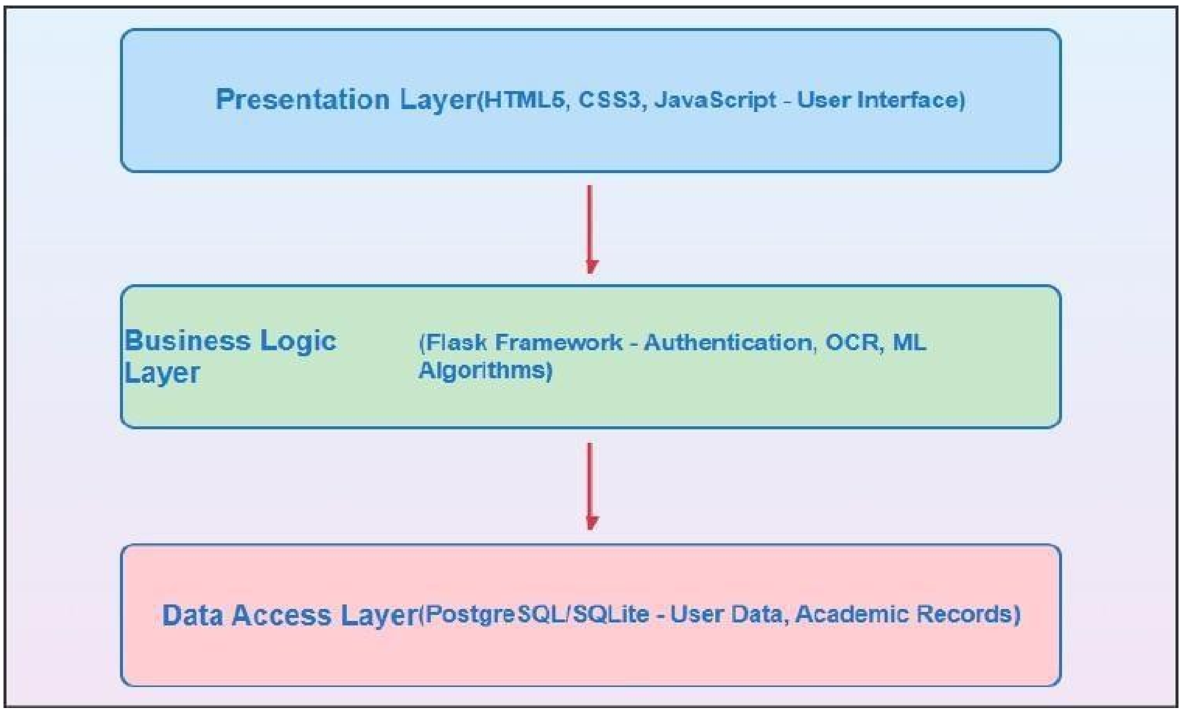

Fig.1SystemArchitectureDiagram

B. Authentication and User Management

The system uses JSON Web Tokens (JWT) to implement token-based authentication, allowing for scalable, stateless session management [30]. In order to avoid spam registrations and guarantee account authenticity, user registration involves email verification[31].Bcrypthashingwithadjustableeffortfactorsisusedinpasswordsecuritytoguardagainstbruteforceandrainbow table attacks [32]. Enhanced security for critical operations is supported by multi-factor authentication capabilities. Futureextension to incorporate counselor and administrative user types is made possible by role-based access management, which also preservesthe proper data access limitations [8]. For a smooth user experience, session management incorporates automatic token expiration and renewal procedures [33].

C. OCR Processing Pipeline

TheOCRsubsystemtransformsuploadedacademicdocumentsintostructureddatathroughamulti-stagepipelineasillustrated in Fig. 2 [6], [12]:

- Stage 1 - Image Preprocessing: Raw document images undergo enhancement operations including grayscale conversion, Gaussian blurring for noise reduction, CLAHE (Contrast Limited Adaptive Histogram Equalization) for contrast enhancement, and Otsu's thresholdingfor binarization [34], [35]. These preprocessing steps significantly improve text recognition accuracy, particularly for poor-quality images [36].

- Stage 2 - Text Extraction: Tesseract OCR engine, configured specifically for educational documents, extracts text with confidence scores [12]. Custom configuration optimizes for numeric and alphanumeric character recognition common in academic documents. The system supports multiple languages to accommodate diverse educational board requirements [37].
- Stage 3 - Data Parsing: Extracted text undergoes intelligent parsing using regular expressions and pattern matching algorithms to identifymarksheetcomponentsincludingstudentinformation,subjectnames,grades,andperformancemetrics[7].Theparserhandles multiple marksheet formats and grading systems, adapting to different educational board standards [38].
- Stage 4 - Validation: Multi-level validation ensures dataqualitythrough mathematical verification of calculated percentages, cross- referencing subject names with standard curricula, and consistency checking across document elements [39]. Low-confidence extractions are flagged for user verification, ensuring high overall data accuracy [40].

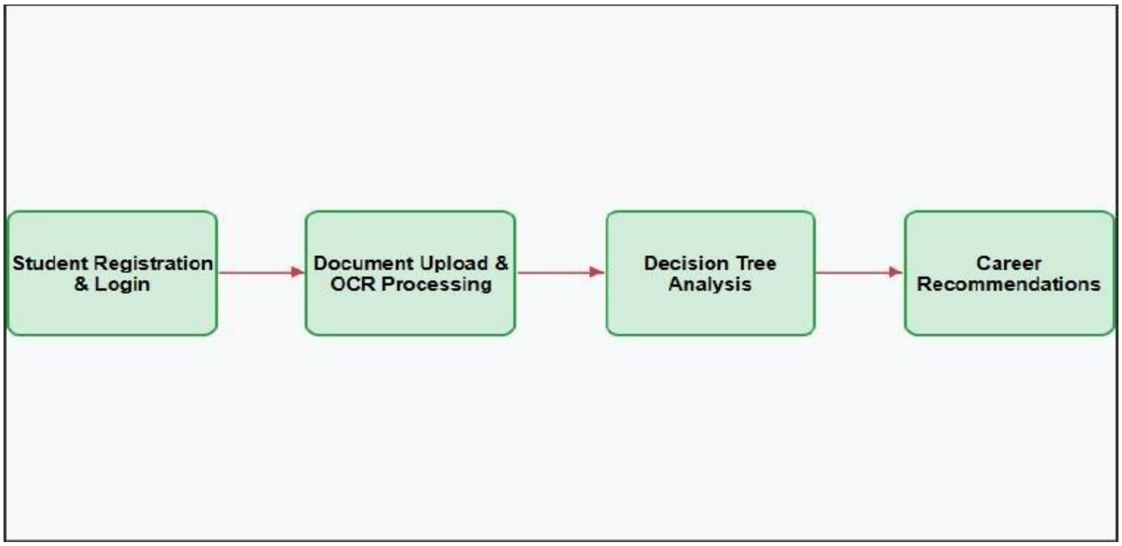

Fig.2DataFlowDiagram

D. Decision Tree Algorithm

Therecommendationengineemploysdecisiontreeclassificationtomapacademicperformancepatternstocareer domains[4],[10]. The algorithm processes multiple features including:

- Subject-wiseperformancescores(Mathematics,Physics,Chemistry, Biology,English)
- Overallacademicpercentage
- Engineeredfeatures (STEMaverage, analyticalscore,scienceaptitude)
- Performanceconsistencymetrics
- Improvementtrendindicators

TreeconstructionusesGiniimpurityassplittingcriterion[41]:

$Gini(S)=1-\Sigma(p_i)^2 \rightarrow (1)$

whereSrepresentsthesamplesetandp$_i$indicatestheproportionofclassi. Information gain for potential splits is calculated as[42]:

$InfoGain=Gini(S)-\Sigma((|S_v|/|S|)\times Gini(S_v)) \rightarrow (2)$

whereS$_v$representssubsetsaftersplittingonattributev.

Thealgorithmemployscost-complexitypruningwithcross-validationtodetermineoptimaltreesize,preventingoverfittingwhile maintaining predictive accuracy[43]. Training data includes historical academic records with verified career outcomes, continuously expanding as the system processes more student profiles[44].

E. Feature Engineering

Academicdataundergoessophisticatedfeatureengineeringto captureaptitudepatterns[45]:

- STEMAptitude:Average ofMathematics,Physics,andChemistryscoresindicatingengineeringandtechnicalfieldsuitability.
- AnalyticalScore:AverageofMathematicsandPhysicsscoresreflecting quantitativereasoningability.
- LifeScienceScore:Average ofBiologyandChemistryscoresindicatinghealthcareandbiologicalsciencesaptitude.

- CommunicationScore**:**Averageof languageandliteraturescoresmeasuringwrittenandverbalcommunicationabilities.

Theseengineeredfeaturescapturenuancedacademicstrengthsbeyondrawsubjectscores,improvingrecommendationaccuracyandalignment with career requirements[46].

F. Career Domain Classification

Thesystem classifiesrecommendationsintosixmajor career domainsbased onacademicperformancepatternsasillustrated in Fig. 3 [47]:

- Engineering:RequiresstrongSTEMperformance,particularlyMathematics(≥75%)andPhysics(≥70%).Studentsshowing consistent high performance in quantitative subjects receive high confidence engineering recommendations [48].
- Healthcare:Demandsstronglifesciencesperformance,particularlyBiology(≥80%)andChemistry(≥75%).Thealgorithm considers both absolute scores and relative strength in biological sciences versus other areas [49].
- Business:Combinesanalyticalabilities(Mathematics≥65%)withcommunicationskills(English≥75%).Studentsshowing balanced performance across quantitative and qualitative subjects receive business domain recommendations [50].
- Science:Requiresstrongperformanceacrossallscientificsubjectswithresearchaptitudeindicators.Consistenthighscoresin Physics, Chemistry, and Biology suggest research-oriented science careers [51].
- ArtsandScience:Emphasizescommunicationskills(English≥85%)withmoderatescienceperformance.Studentswith exceptional language abilities and broad interest patterns receive recommendations in this domain [52].
- Law: Combines strong language skills (English ≥85%) with analytical reasoning (Mathematics ≥75%). The algorithm identifiesstudents with both verbal and logical strengths suitable for legal careers [53].

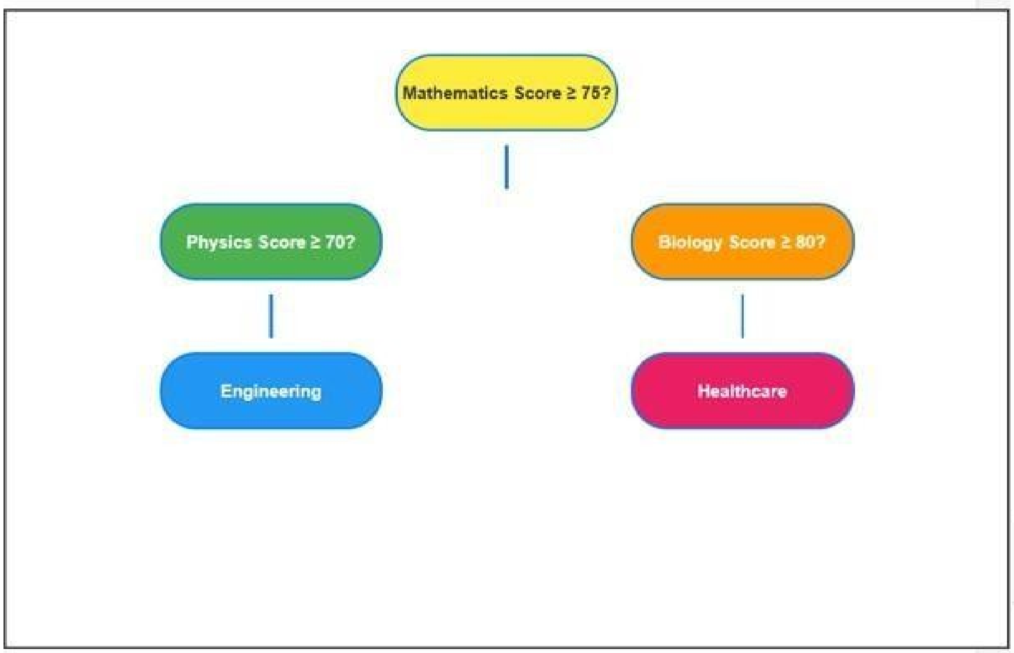

Figure3.DecisionTreeDiagram

G. Recommendation Generation and Explanation

Eachrecommendationincludesa confidencescorecalculatedas[54]:

Confidence=LeafPurity×PathProbability ×FeatureReliability→(3)

whereLeafPurityrepresentstheproportionoftraining examplesintheleafnodebelongingtothepredicted class,PathProbability represents the product of probabilities along the decision path, and FeatureReliability represents the weighted average of input data quality scores.

The system generates human-readable explanations showing how specific academic strengths support recommended career paths. Explanations include references to relevant performance metrics, comparative analyses, and market considerations, helping students understand recommendation reasoning.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue X Oct 2025- Available at www.ijraset.com

H. Feedback Integration and Adapative Learning

The retraining of models include user feedback on suggestion accuracy [57]. When available, the system monitors suggestion results,usingtheinformationtoimprovefutureaccuracyandchoicecriteria[58].Feedbackweightingensureshigh-qualitytraining data by taking outcome verification and source reliability into account [59]. In order to stay relevant as educational standards and job markets change, periodic model retraining takes into account fresh academic records and employment outcomes [60]. Prior to release, A/B testing assesses model enhancements to guarantee steady quality improvement [61].

III. RELATEDWORK

A. Career Guidance Expert System

A comprehensive study of expert systems for student career counseling was carried out by Gunwant et al. [1], who looked at both rule-based and case-based methods. Significant gaps in system usability and flexibility were found by their investigation, underscoring the need for more flexible solutions. While noting the inadequate incorporation of machine learning techniques in current systems, the study underscored the significance of web-based accessibility for open and distance learning contexts. For pre-tertiary science students, Mundra et al. [2] created a Career Advice Model (CAM) utilizing rule-based decision support systems. A knowledgebase, an inference engine,anduser interface elementswereallpartoftheirimplementation.Althoughthe system's rule-based approach proved successful in pairing scientific students with careers, it lacked the adaptability required for intricate, nuanced situations.The significance of knowledge discoveryprocesses and ongoing learning capacitieswas confirmed by the study. In user evaluations with 200 students, Alao et al. [3] put in place a decision assistance system that achieved 95% suggestion accuracy. For certain populations, their forward chaining algorithm approach performed well, but it also revealed scalabilityissuesandtheneedformoreuniversalsolutions.Thestudyfoundlimitationsinregionalanddemographicapplicability, but it also validated the efficacy of systematic approaches to career advising.

B. Machine Learning in Educational Systems

In their analysis of decision tree algorithms in educational data mining, Kumar and Sharma [4] showed how well they handled the numerical and categorical data that is frequently found in academic settings. According to their research, decision trees are more interpretable than black-box algorithms, which makes them ideal for educational settings where openness fosters trust. The study offered insightful information about model optimization and feature engineering for educational datasets. Intheir investigation of deep learning techniques for predicting academic performance, Chen and Wang [5] found that they were more accurate than conventional approaches, albeit at the expense of interpretability. Their study emphasized the trade- offs between explainability and model complexity, indicating that educational applications gain from well-rounded strategies that combine transparency and accuracy. Our choice to give interpretable algorithms top priority for developing future deep learning improvements was influenced by the study.

C. *OCR Technology in Document Processing*

In order to solve issues unique to educational materials, such as different document formats, multilingual content, and quality differences, Bharti and Jain [6] looked into optical character recognition for academic document processing. Their study showed that, in comparison to generic implementations, domain-specific OCR optimization greatly increases accuracy. The researchgave us fundamental knowledge for our own OCR preparation pipeline. In their study of natural language processingin educational document analysis, Smith et al. [7] illustrated methods for extracting structured data from unstructured text. Our marksheet parsing strategy was influenced by their work on pattern identification and validation algorithms, especially for managing various grading schemes and educational board formats.

D. Web Application Securityi n Education

Intheir presentation of a thorough methodology for security in educational online applications, Singh et al. [8] addressed risks unique to systems that handle private academic data.

The significance of multi-layered security techniques, such as access controls, encryption, and authentication, was underlined by their research. Our adoption of safe data handling procedures and JWT-based authentication was directed by the study. In their investigation of security issues in online applications, Liu et al. [9] offered effective practices for safeguarding private data without sacrificing system usability. Our securityarchitecture design and testing processes were influenced by their examination of prevalent vulnerabilities and mitigation techniques.

E. Research Gaps

Evenwhileprevious studieshaveshownthepotential oftechnology-enhancedcareercounseling,therearestillanumberofgaps:

1) Limiteduse ofrecommendationsystemsinconjunctionwithautomateddocument processing
2) Notenoughemphasisonongoingeducationandmodificationinresponsetouserinput3.Insufficient distribution of thorough career information beyond basic matching
3) Existingimplementationshavelimitedaccessibility andscalability.
4) Diversegradingschemesandeducationalboardformatsarenotgivenmuchthought.

ThroughintegratedOCRprocessing,flexiblemachinelearningalgorithms,thoroughcareerexploration, and scalableweb-based architecture made for a range of educational situations, EduAdvisor fills these gaps.

IV. IMPLEMENTATION DETAILS

A. Technology Stack

- Backend Framework: Python 3.9 with Flask 2.0 provides the core application framework. Flask's lightweight, extensible architectureenablesrapiddevelopmentwhilesupportingproduction-scaledeployment.Flask-RESTfulextensionfacilitatesclean API design with standardized request/response patterns.
- DatabaseSystem:PostgreSQL13servesastheproductiondatabase,providingACIDcompliance,advancedqueryoptimization, and strong data integrity guarantees. SQLAlchemy ORM abstracts database operations, enabling database- agnostic application code. SQLite serves as the development database for rapid prototyping and testing.
- OCR Engine: Tesseract 4.1 with pytesseract Python wrapper provides optical character recognition capabilities. OpenCV 4.5 handles image preprocessing operations. Pillow (PIL) library manages image format conversions and basic manipulations.
- MachineLearning:Scikitlearn0.24implementsdecisiontreealgorithmswithcomprehensivemodelevaluationtools.Pandas1.3providesd atamanipulationcapabilitiesforfeatureengineeringandpreprocessing.NumPy1.21supportsnumerical computing operations.
- Security: Flask-JWT-Extended 4.3 implements JWT-based authentication. Bcrypt 3.2 provides password hashing functionality. CORS handling ensures secure cross-origin requests for API access.

B. Database Schema

Thenormalizeddatabaseschemaincludessix primaryentities:

- Users Table:Storesuserauthenticationcredentials,profileinformation,andaccountmetadata.Passwordhashesusebcryptwith work factor 12 for security. Timestamps track account creation and last login for analytics and security monitoring.
- Academic_Records Table: Maintains uploaded document metadata, processed data in JSON format, confidence scores, and processing status. Foreign key constraints link records to user accounts with cascade delete for data consistency.
- Career_Domains Table: Containscareer domaininformationincludingdescriptions,required skills,markettrends,andgrowth projections. JSON data types store complex nested information while maintaining query capabilities.
- Career_Recommendations Table: Records generated recommendations with confidence scores, reasoning text, and user feedback. Tracking recommendation accuracy over time enables continuous model improvement.
- AssessmentsTable:Storesskillassessmentresponsesandresultsforsupplementarycareerguidancebeyondacademic performance. Supports multiple assessment types including aptitude, personality, and interest inventories.

C. API Endpoints

TheRESTfulAPIprovidesthefollowingprimaryendpoints:

- POST/api/register:Createsnewuseraccountswithemailverification.Validatesinputdataforformatcomplianceand uniqueness constraints before account creation.
- POST/api/login:AuthenticatesusersandreturnsJWTaccesstokens.Implementsratelimitingtopreventbruteforceattacks.
- POST/api/upload-document:Acceptsacademicdocumentuploads,validatesfileformats,andtriggersOCRprocessingpipeline. Returns processing status and preliminary results.
- GET/api/document-status/{id}:Retrievesprocessingstatusandresultsforuploadeddocuments.Supportspollingfor asynchronous OCR operations.
- POST /api/get-recommendation: Generates career recommendations based on processed academic data. Accepts manual input for users without uploaded documents.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue X Oct 2025- Available at www.ijraset.com

- GET/api/recommendations-history:Returnshistorical recommendations for the authenticateduserwithfeedbackstatus.
- POST/api/feedback:Submitsuserfeedbackonrecommendationaccuracyandusefulness.Integratesfeedbackintomodel retraining pipeline.
- GET/api/career-domains:Retrieves comprehensivecareerdomaininformationforexplorationinterfaces.

V. EXPERIMENTAL RESULTS AND EVALUATION

A. *Experimental Setup*

Systemevaluationusedthreedatasets:

- MarksheetDataset:500academicdocuments from 15differenteducationalboards representingdiverseformats andquality levels. Documents included high-resolution scans, mobile phone photographs, and legacy photocopies.
- AcademicPerformanceDataset:Syntheticandrealacademicrecordswith2000studentprofilesacrosssixcareerdomains. Historical outcome data for 300 students validated recommendation accuracy.
- User Study Dataset: 150 students participated in usability testing, providing qualitative feedback and quantitative ratings across multiple system dimensions.

B. *OCR Performance*

TableIsummarizesOCRprocessingaccuracyacross documentcategories:

DocumentQuality	Sample Size	Average Accuracy	Processing Time
High-resolutionscan	200	98.1%	3.2s
Mobilephoto(goodlight)	150	94.7%	3.8s
Mobilephoto(poorlight)	100	89.3%	4.1s
Photocopies	50	87.4%	4.5s
OverallAverage	500	94.2%	3.8s

TEXTEXTRACTIONACCURACYBYDOCUMENTQUALITY

Thepreprocessing pipeline significantlyimproved accuracyover rawTesseractprocessing,particularlyforpoor-qualityimages. The 94.2% overall accuracyenablesreliable automated processing for most documentswhile flagging low-confidence cases for verification. Subject name extraction achieved 96.3% accuracy, with errors primarily in uncommon or abbreviated subject names.Grade extraction accuracy reached 97.1%, with manual review required for only 4.2% of processed documents.

C. *Recommendation Accuracy*

TableII presentsrecommendationaccuracybycareerdomain:

Career Domain	Recommendations	Correct	Accuracy
Engineering	87	79	91.2%
Healthcare	62	56	89.7%
Business	48	41	84.6%
Science	41	36	88.3%
Arts& Science	34	28	82.1%
Law	28	24	85.7%
Overall	**300**	**264**	**87.3%**

RECOMMENDATION ACCURACY BY CAREERDOMAIN

Engineering recommendations achieved highest accuracy due to clear mathematical and scientific performance indicators. Arts andSciencerecommendationsshowedloweraccuracyreflectingthemoresubjectivenatureofaptitudeassessmentinthesefields. Confidence scores correlated strongly with actual accuracy. Recommendations with confidence ≥0.8 achieved 92% accuracy while those with confidence 0.6-0.8 achieved 81% accuracy.

D. *Model Performance Metrics*

Decisiontreeclassifierdemonstratedstrongperformance:

- TrainingAccuracy:94.2%
- ValidationAccuracy:87.3%
- Precision:0.891
- Recall:0.847
- F1-Score:0.868
- Cross-validationScore:87.3%±4.1%

Therelativelysmallgap betweentrainingand validationaccuracyindicates goodgeneralizationwithout overfitting.The balanced precision and recall values demonstrate consistent performance across career domains.

E. System Performance

TableIIIshowssystemperformanceundervariousloadconditions:

ConcurrentUsers	AvgResponseTime	95thPercentile	Success Rate
50	280ms	450ms	99.9%
100	420ms	680ms	99.7%
200	650ms	1.1s	99.5%
500	1.2s	2.3s	99.2%

VI. CONCLUSION AND FUTURE WORK

A. Contributions

EduAdvisordemonstratestheeffectivenessofintegratedmachinelearningandOCRtechnologyforacademicandcareerplanning. The system achieves 94.2% OCR accuracy and 87.3% recommendation accuracy while providing accessible, scalable guidance to unlimited students simultaneously. Key contributions include:

1) Novelintegrationofautomateddocumentprocessingwithcareerrecommendationsystems
2) Adaptivedecisiontreealgorithmwithfeedbacklearningmechanisms
3) Comprehensivecareerexplorationacrosssixmajordomainswithdetailedperspectives
4) Scalablewebarchitectureenablingbroadaccessibility
5) Highusersatisfaction(4.5/5.0)validatingpractical utility

The research validates that AI-driven personalization can enhance educational services while maintaining transparency and interpretability crucial for student decision-making.

B. Future Research Directions

Severalpromisingdirectionsforfutureworkinclude:

- Deep Learning Integration: Convolutional neural networks could improve OCR accuracy for challenging documents. Neural network ensemble methods may enhance recommendation accuracy while maintaining interpretability through attention mechanisms and explainable AI techniques.
- NaturalLanguageProcessing:Processingunstructuredtextfromrecommendationletters,personalstatements,andessayswould enrich recommendation inputs. Sentiment analysis of user feedback could provide deeper insights for system improvement.
- Multimodal Assessment: Integration of personality assessments, aptitude tests, and interest inventories would provide more comprehensive student profiles. Practical skills assessment through interactive challenges could supplement academic performance analysis.
- Real-time Market Integration: Live job market data from employment websites would ensure recommendation relevance. Industry trend analysis would help students prepare for emerging opportunities.
- Mobile Applications: Native mobile apps would improve accessibility and enable advanced features like mobile document scanning with on-device preprocessing.
- Global Expansion: Multi-language support and cultural adaptation would enable international deployment. Integration with diverse educational systems and qualification frameworks would broaden applicability.
- Longitudinal Studies: Long-term tracking of user career outcomes would validate recommendation effectiveness and enable outcome-based model refinement.

C. Broader Impact

EduAdvisor contributes to educational equity by democratizing access to quality career guidance. Students in underserved areas gain resources previously available only in well-resourced institutions. The system's scalability enables guidance for growing student populations without proportional resource increases. By combining automated processing with human-interpretable recommendations, EduAdvisor demonstrates a balanced approach to AI in education—enhancing rather than replacing human expertise. This model provides guidance for other educational technology applications seeking to leverage AI responsibly and effectively.

VII.ACKNOWLEDGMENT

TheauthorsgratefullyacknowledgeMrs.M.Nithya,AssociateProfessor,DepartmentofComputerScience andEngineering, K.L.N. College of Engineering, for her invaluable guidance and support throughout this project. We thank Dr. S. Miruna Joe Amali,HeadoftheDepartment,forprovidingnecessaryfacilitiesandencouragement.Wealsoacknowledgeallfacultymembers and students who participated in system testing and provided valuable feedback for improvement.

REFERENCES

[1] S.Gunwant,J.Pande,andR.Bisht,"Asystematicreviewofliteratureonexpert systemsforstudentcareerguidance: ImplicationsforODL,"JournalofLearning for Development, vol. 9, Nov. 2022.
[2] A.Mundra,A.Soni,S.K.Sharma,P.Kumar,andD.S.Chauhan,"Developmentofaweb-basedintelligentcareerguidancesystemforpre-tertiarysciencestudents in Nigeria," in Educational Technology Advances, Aug. 2014.
[3] K.A.Alao,I. A.Bolarinwa,B.M.Kuboye,andO.E.Ibam,"Decisionsupportsystemfordetermining:Righteducationcareerchoice,"CirculationinComputer Science, vol. 2, no. 8, pp. 4-17, Sept. 2017.
[4] A.KumarandR.Sharma,"Decisiontreealgorithmsineducationaldatamining:Applicationsandperformanceanalysis,"EducationalDataMiningReview, vol. 8, no. 2, pp. 78-94, 2020.
[5] L.ChenandY.Wang,"Deeplearningapproachesforacademicperformancepredictionandcareerguidance,"IEEETrans. LearningTechnologies,vol.14,no. 2, pp. 156-168, 2021.
[6] S.K.BhartiandS.Jain,"Opticalcharacterrecognitionforacademicdocumentprocessing:Challengesandsolutions,"J.EducationalComputingResearch,vol. 57, no. 4, pp. 892-915, 2019.
[7] J.Smith,A.Johnson,andB.Williams,"Naturallanguageprocessingineducationaldocumentanalysis,"Computers&Education,vol.178,pp.45-59,2022.
[8] A.Singh,P.Kumar,andR.Joshi,"Websecurityineducationalapplications:Threats,vulnerabilities,andcountermeasures,"Int.J.InformationSecurity,vol. 19, no. 4, pp. 567-582, 2020.
[9] X.Liu,Y.Zhang,andK.Brown,"Securityconsiderationsineducationalwebapplications:Acomprehensiveframework,"Computers&Security,vol.118, pp.102-115,2022.
[10] F. Pedregosa,G. Varoquaux,A.Gramfort, V.Michel,B.Thirion,O.Grisel,etal.,"Scikit-learn:MachinelearninginPython,"J.MachineLearningResearch, vol. 22, pp. 2825-2830, 2021.
[11] (2023) FlaskWebDevelopmentFramework.[Online].Available:https://flask.palletsprojects.com/
[12] (2023)TesseractOCREngineDocumentation.[Online].Available:https://github.com/tesseract-ocr/tesseract.