0011110001110001000100100101010101010101010101010101010101010101010101010101010101010101010101010101

الجامعة العربية الامريكية

فلسطين

توقع التخصصات الجامعية لطلاب البكالوريوس باستخدام خوارزميات الذكاء الاصطناعي: نحو تطوير نظام إرشاد ذكي للطلاب الجدد

رامي حدرب
عميد القبول والتسجيل
الجامعة العربية الأمريكية - فلسطين

طلبتنا الأعزاء في الثانوية العامة
مستمرون معكم في التسجيل والأيام الإرشادية
من السبت حتى الخميس من كل أسبوع
4:00 PM - 8:00 AM
في حرمي الجامعة جنين ورام الله
طبيبة
علمٌ بلا حُدود

الجامعة العربية الأمريكية
ARAB AMERICAN UNIVERSITY
دكتورة أسنان

الطلبة الجدد
مهندس

استعلامات الطلبة الجدد

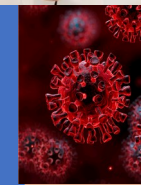
علمٌ بلا حُدود

الجامعة العربية الأمريكية
لنصنع مستقبلاً مشرقاً
مهندسة

الجامعة العربية الأمريكية
Done

ممرضة

سجل الآن
بكالوريوس علوم الواقع الافتراضي



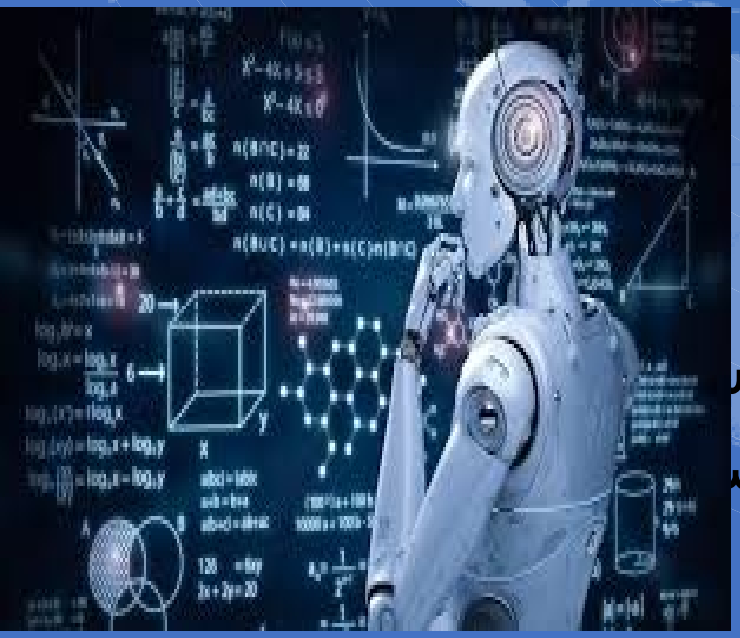
مقدمة

- تهدف مؤسسات التعليم العالي إلى أتمتة العمليات مثل القبول والتسجيل والإرشاد الأكاديمي:

✓اختيار الطالب لتخصصه له آثار كبيرة على مسيرته المهنية المستقبلية.

✓يجب على الجامعات الانتقال من الأنظمة التقليدية إلى الأنظمة الذكية باستخدام الذكاء الاصطناعي وتعلم الآلة.

✓يظهر تعلم الآلة --Machine Learning أنماطًا مخفية في البيانات، مما يعزز عملية اتخاذ القرار.



- **الفوائد الرئيسية:**
- تحسين معدلات التخرج وتقليل معدلات التسرب.
- أنظمة إرشاد ذكية لتوجيه الطلاب في اختيار التخصصات المنا
- دعم النمو الاقتصادي من خلال مواءمة التعليم مع احتياجات س

نظرة عامة على الدراسة

- تقترح هذه الدراسة طريقة ذكية لمساعدة الطلاب الجدد في الجامعة العربية الأمريكية على اختيار تخصصهم الدراسي (التخصص الأكاديمي).
  الميزات الرئيسية:
- استخدام نماذج تعلم الآلة على بيانات طلاب الجامعة العربية الأمريكية.
- تطوير نظام إرشاد يتضمن مراحل: إعداد البيانات، عرض البيانات، اختيار نموذج تعلم الاله، والتطبيق.
- **نظرة عامة على النظام**
  الهدف:
  تحسين كفاءة التنبؤ من خلال تحليل إحدى عشرة/ثمانية خصائص أساسية.
- **مجموعة البيانات:**
  8620 سجل يحتوي على تفاصيل الطلاب.
- **المتغير المستهدف:**
  اختيار تخصص الدراسة الجامعية بناءً على السمات الأكاديمية والديموغرافية والطبية.

Related Research

- Several studies have focused on machine learning-based student major selection:
- King Abdulaziz University (KAU): KNN algorithm with 70.83% accuracy on preparatory students (2017).
- Saudi University: Data mining techniques predicting early academic success using General Aptitude Test, Scholastic Achievement Test, and High School Grades.
- Al-Azhar University: Adaptive recommendation system for selecting engineering departments, achieving 82.57% accuracy with Random Forest.
- Decision Support Systems (DSS): Naïve Bayes (NB), Random Forest (RF), and Decision Trees used to recommend suitable majors, with RF achieving accuracy of up to 97.87%.
- Various Case Studies: Application of ANN, SVM, and Logistic Regression models with accuracies ranging from 80% to 97.7%.

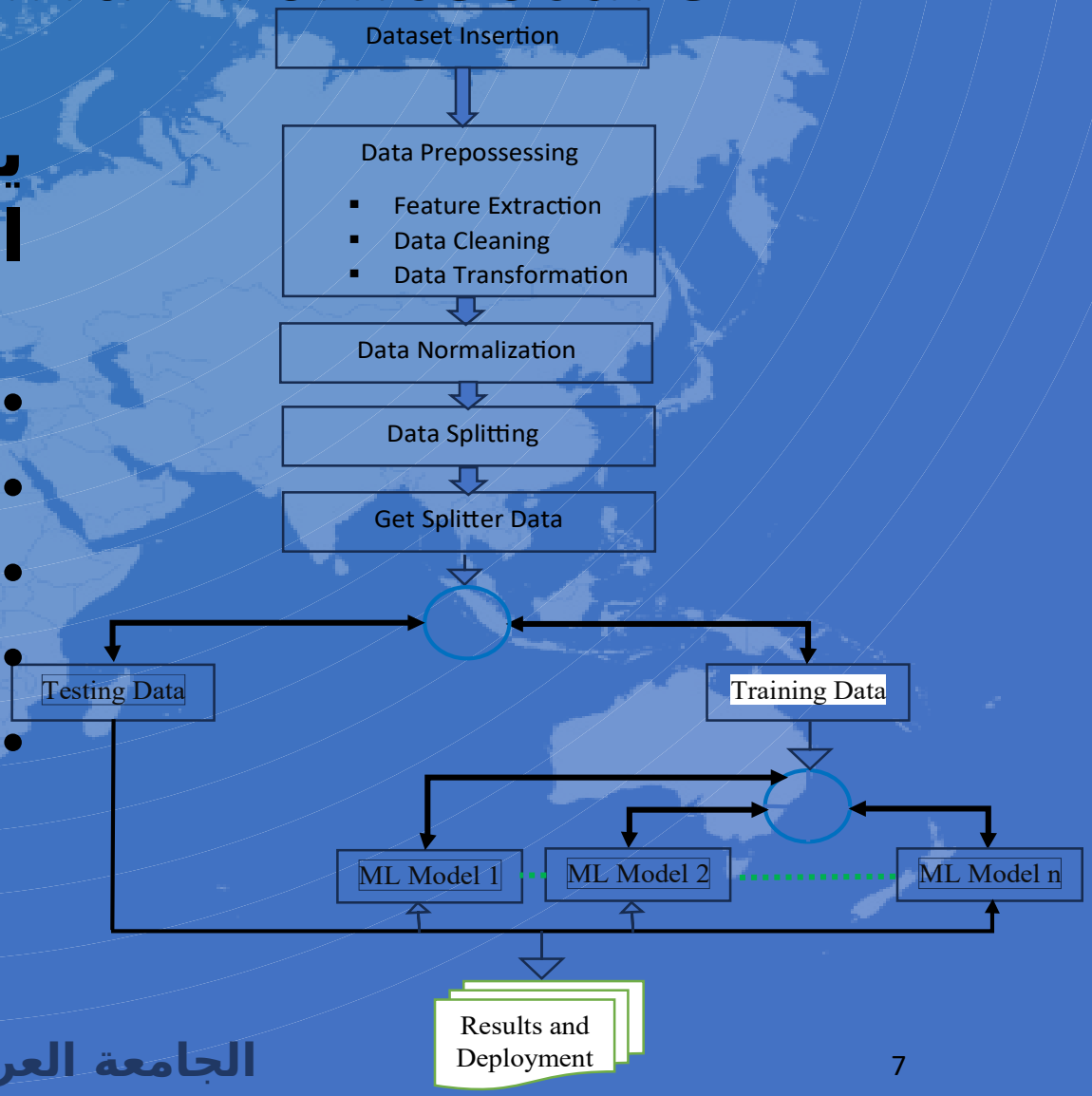
وصف مجموعة البيانات

- **نظرة عامة على مجموعة البيانات**
- **المصدر:** طلبات القبول في الجامعة العربية الأمريكية (2021–2024).
- **عدد السجلات:** 8620 سجلًا للطلاب، 8/11 خاصية مدخلة، ومتغير مستهدف واحد (التخصص الدراسي).
- **تشمل الخصائص:**
- **البيانات الأكاديمية للطلاب:** فرع المدرسة الثانوية (مثل العلمي، الأدبي)، المعدل.
- **البيانات الديموغرافية:** الجنس، مدينة الإقامة.
- **المستوى التعليمي للوالدين.**
- **الحالة الصحية للطلاب.**
- **خياران**
- **تفضيلات الطلاب:** الكلية المفضلة/أفضل 3 خيارات دراسية.
- **المصدر:** طلبات الالتحاق الجامعة العربية الأمريكية (2021–2024).

Model Methodology and Architecture

يتكون نظام الإرشاد من المراحل التالية

- إعداد البيانات
- عرض البيانات
- اختيار النموذج تعلم الاله
- تطبيق النموذج
- تضمن كل مرحلة دقة وقابلية استخدام نظام الإرشاد.



النماذج المستخدمة

- Machine Learning Models:
  1. Logistic Regression with oversampling.
  2. Decision Trees: Captures and classifies occurrences effectively.
  3. Random Forest: Reduces overfitting with ensemble methods.
  4. Support Vector Machines (SVM): Effective for classification.
  5. K-Nearest Neighbors (KNN): Categorizes large observations accurately.
  6. Naive Bayes: Handles evenly distributed class distributions.

- **الهدف:**
- مساعدة الطلاب في اختيار تخصصهم بناءً على توصيات مستنيرة.
- تحسين كفاءة الإرشاد لعمادة القبول والتسجيل.

Model Performance Metrics (First Option) after Oversampling

Model	Accuracy	Precision	F1 Score	Performance Notes
Logistic Regression	61.4%	61.2%	59.7%	Moderate improvement in linear data
Decision Tree	99.6%	99.2%	99.1%	Nearly flawless performance
Random Forest	99.6%	99.3%	99.2%	Exceptional performance, robust classification
Support Vector Machine (SVM)	78.2%	78.3%	77.4%	Improved accuracy after oversampling
K-Nearest Neighbors (KNN)	98.3%	98.3%	98.3%	High accuracy, grouped observations effectively
Naive Bayes	75.2%	77.2%	73.6%	Moderate improvement in patterns across classes

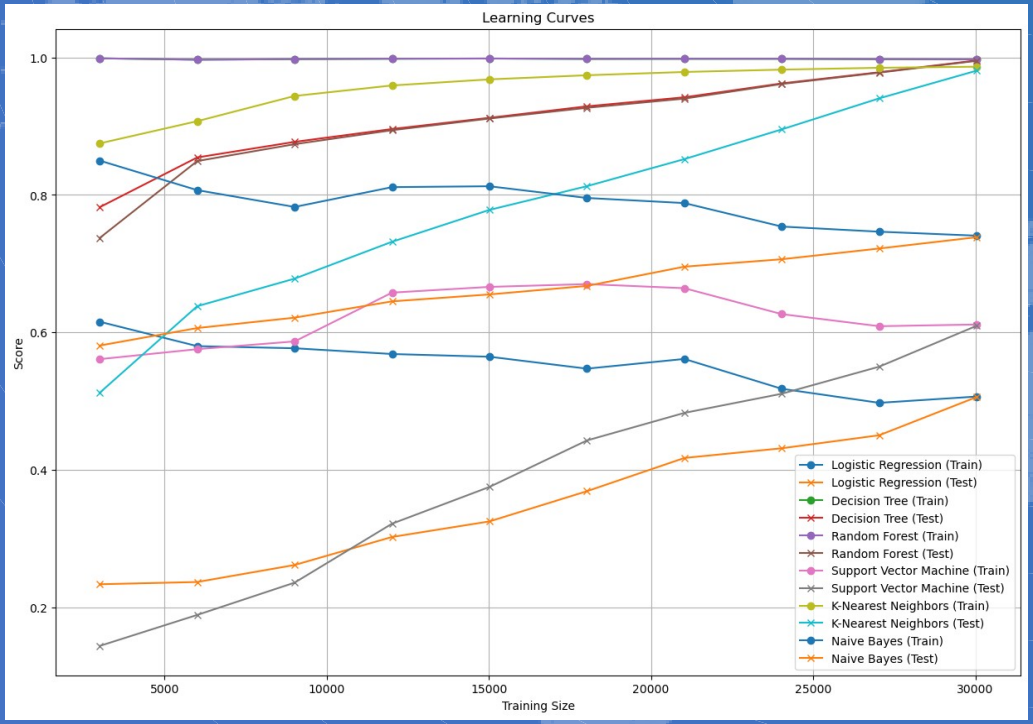
Feature Importance Analysis

Feature	RF Importance	GB Importance	Combined Importance	Notes
FIRST_DESIRE_MAJOR	0.413	0.533	0.533	Most important feature
FACULTY_NAME_S	0.317	0.337	0.337	Second most important feature
SCHOOL_AVG	0.083	0.027	0.083	Significant for academic evaluation
SCHOOL_BRANCH	0.049	0.003	0.049	Moderate relevance
GENDER	0.010	0.039	0.039	Least important feature

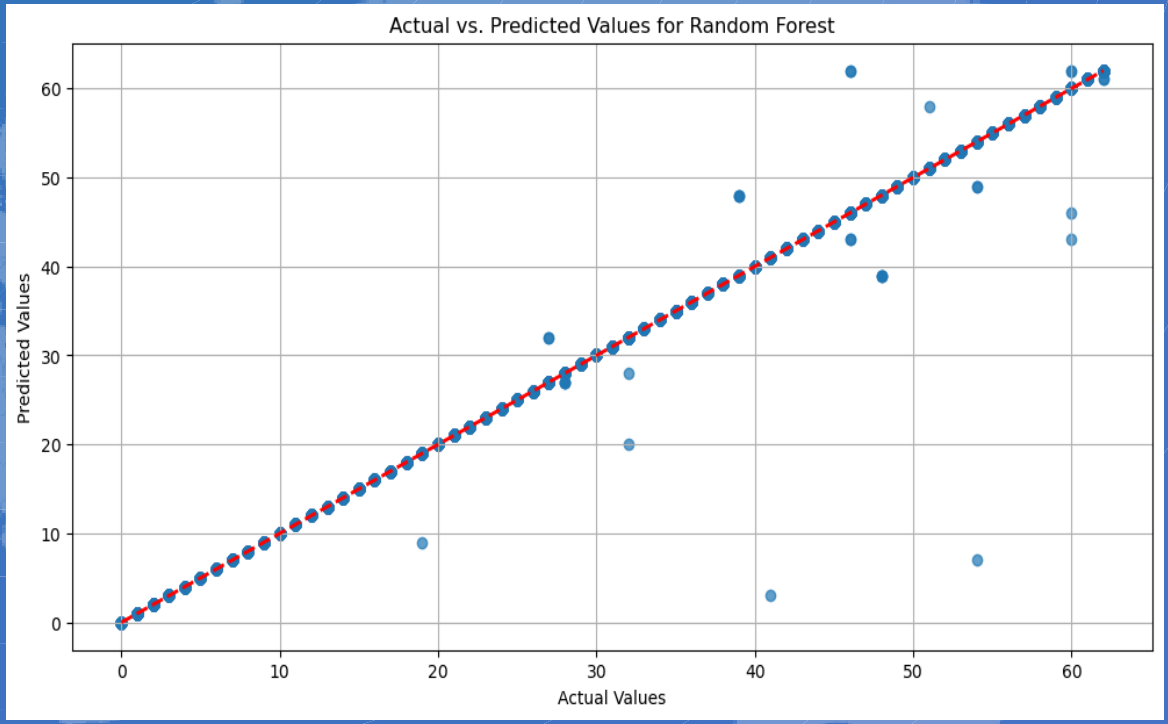
Key Insights:

- Focus on FIRST_DESIRE_MAJOR and FACULTY_NAME_S for major predictions.
- Simplify models by excluding less significant features (e.g., GENDER).
- Random Forest and Decision Tree are optimal for advising systems.
- Oversampling effectively balances data for improved model accuracy.

منحنى التعلم للنماذج



القيم الفعلية مقابل المتوقعة



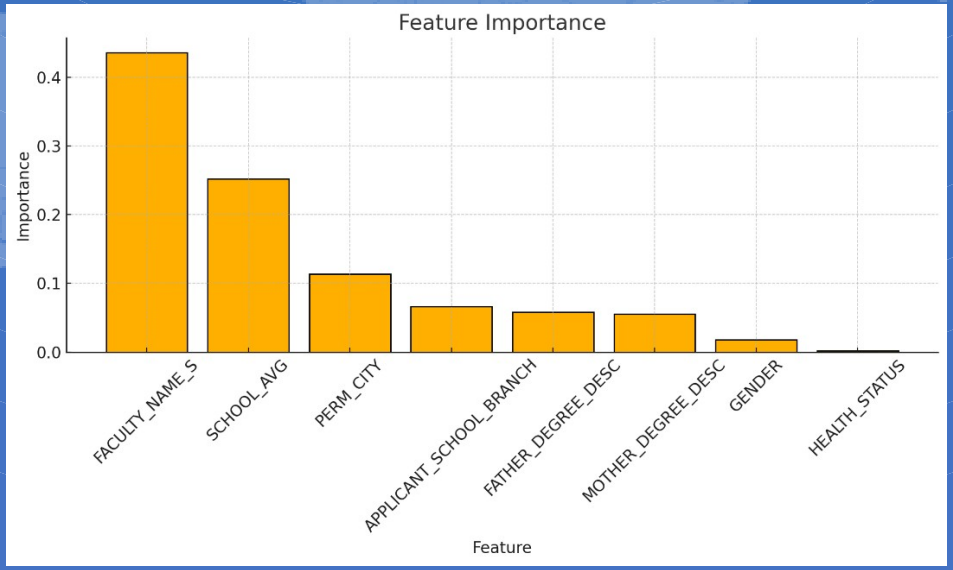
Performance Comparison of ML Models (Second Option)

- Key Input Features:
  - Preferred Faculty Name
  - School Branch (text)
  - School Average (float)
  - Mother and Father Degree Descriptions
  - Permanent Residence City (PERM CITY)
  - Gender
  - Health Status (special medical requirements)

Model	Accuracy	F1 Score	Precision
K-Nearest Neighbors	0.78	0.76	0.75
Random Forest	0.85	0.83	0.82
Decision Tree	0.80	0.78	0.77
Support Vector Machine	0.82	0.81	0.79
Logistic Regression	0.81	0.80	0.79
Naive Bayes	0.75	0.73	0.72
Neural Network	0.84	0.82	0.81

Feature Importance

- Feature importance chart illustrating the contribution of each feature to predicting the study major.
- Random Forest identified Preferred Faculty Name, School Average, and School Branch as critical factors.
- Random Forest assigns the highest importance to these features.



Performance Improvement

- Key Optimizations for Improved Model Performance:
  - Hyperparameter tuning for Random Forest, SVM, and Neural Network.
  - Feature engineering to enhance model interpretability.
  - Data balancing using SMOTE (Synthetic Minority Over-sampling Technique).
  - Model ensemble with Gradient Boosting and Stacking.
- Stacking ensemble combining Random Forest, SVM, KNN, and a Logistic Regression meta-model.
  - Five-fold cross-validation for model generalization.

Performance Metrics for Individual Models and Stacking

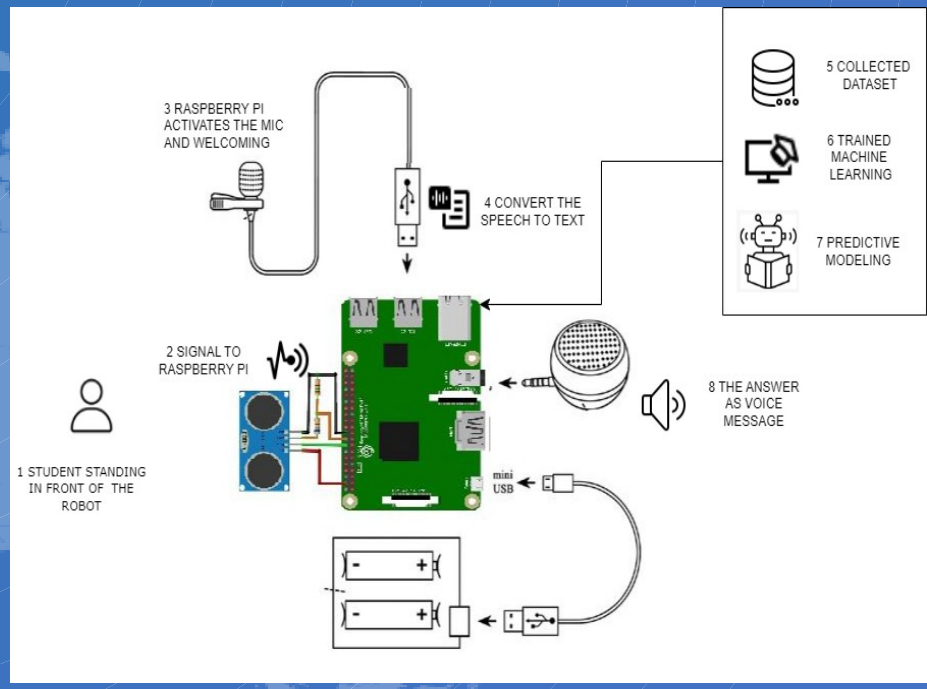
- Integrating SMOTE and Stacking yielded significant performance improvements:
  - Enhanced results across Random Forest, SVM, and KNN base models.
  - Meta-model (Logistic Regression) improved ensemble performance.
  - SMOTE balanced minority classes, reducing class imbalance effects.

Model	Test Accuracy	F1 Score	Precision	Recall
Random Forest	0.88	0.87	0.88	0.87
SVM	0.83	0.82	0.84	0.81
Neural Network	0.85	0.82	0.83	0.81
KNN	0.78	0.77	0.78	0.77
SMOTE + Stacking	0.9	0.89	0.9	0.89

نظرة عامة على نموذج نظام الإرشاد الذكي

- قمنا بإنشاء نموذج أولي للروبوت باستخدام خوارزمية (Random Forest)) كنموذج تنبؤ لنشر نظام إرشاد ذكي لاختيار تخصص الطلاب الجدد.
  تم تصميم النظام بعناية لتلبية المتطلبات التالية:
- مستشعر فوق صوتي للكشف عن الحضور.
- ميكروفون ومكبر صوت لأوامر الصوت وإخراج الكلام.
- تكامل الذكاء الاصطناعي ومعالجة اللغة الطبيعية (NLP) باستخدام TensorFlow Lite، وNLTK، وGoogle Speech Recognition
- مكتبة تحويل النص إلى كلام للمحادثة الطبيعية.

Hardware Components



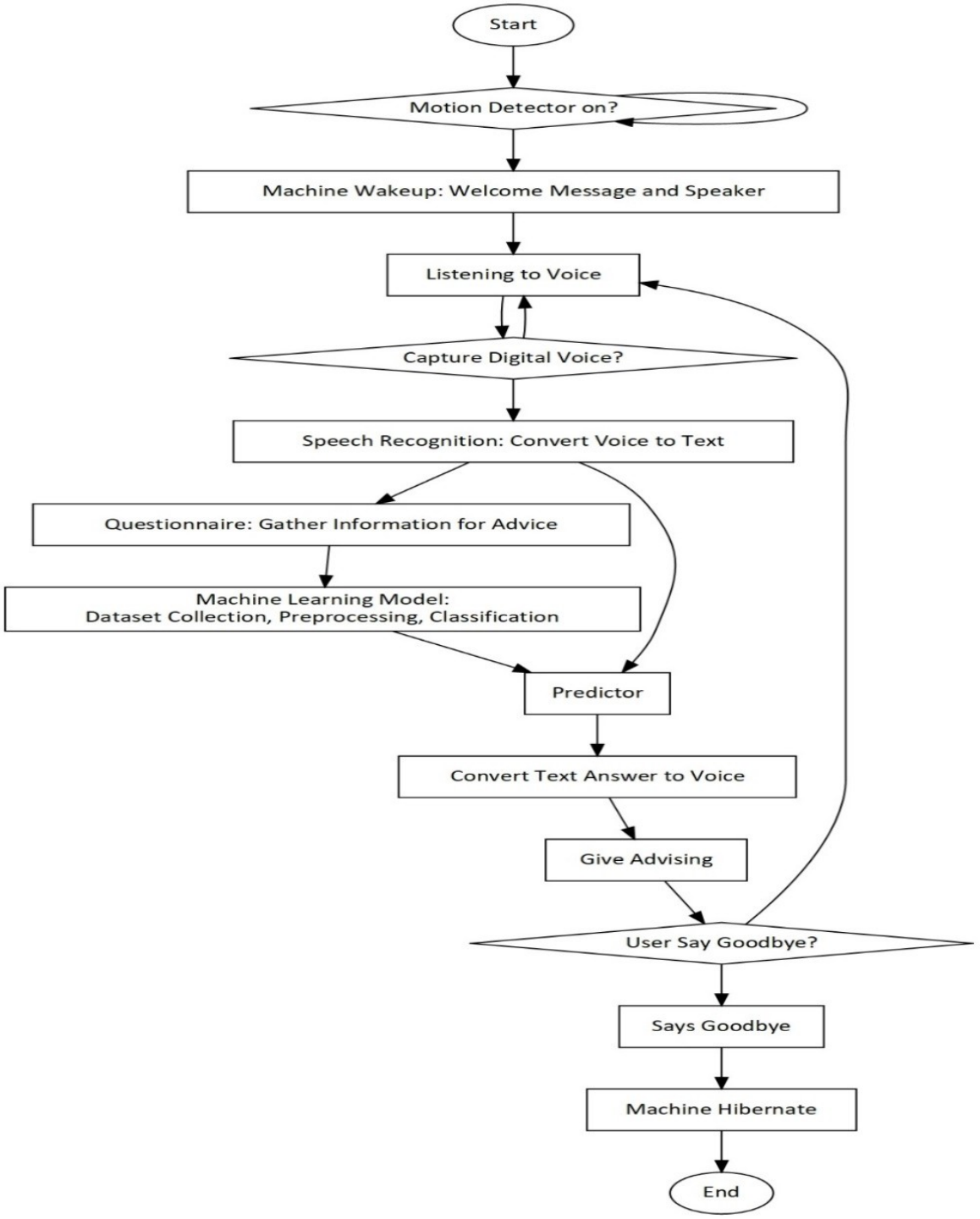
- Aurduino platform with Raspbian OS.
- Ultrasonic sensor for detecting student presence.
- Microphone for voice input and speaker.

Software Integration

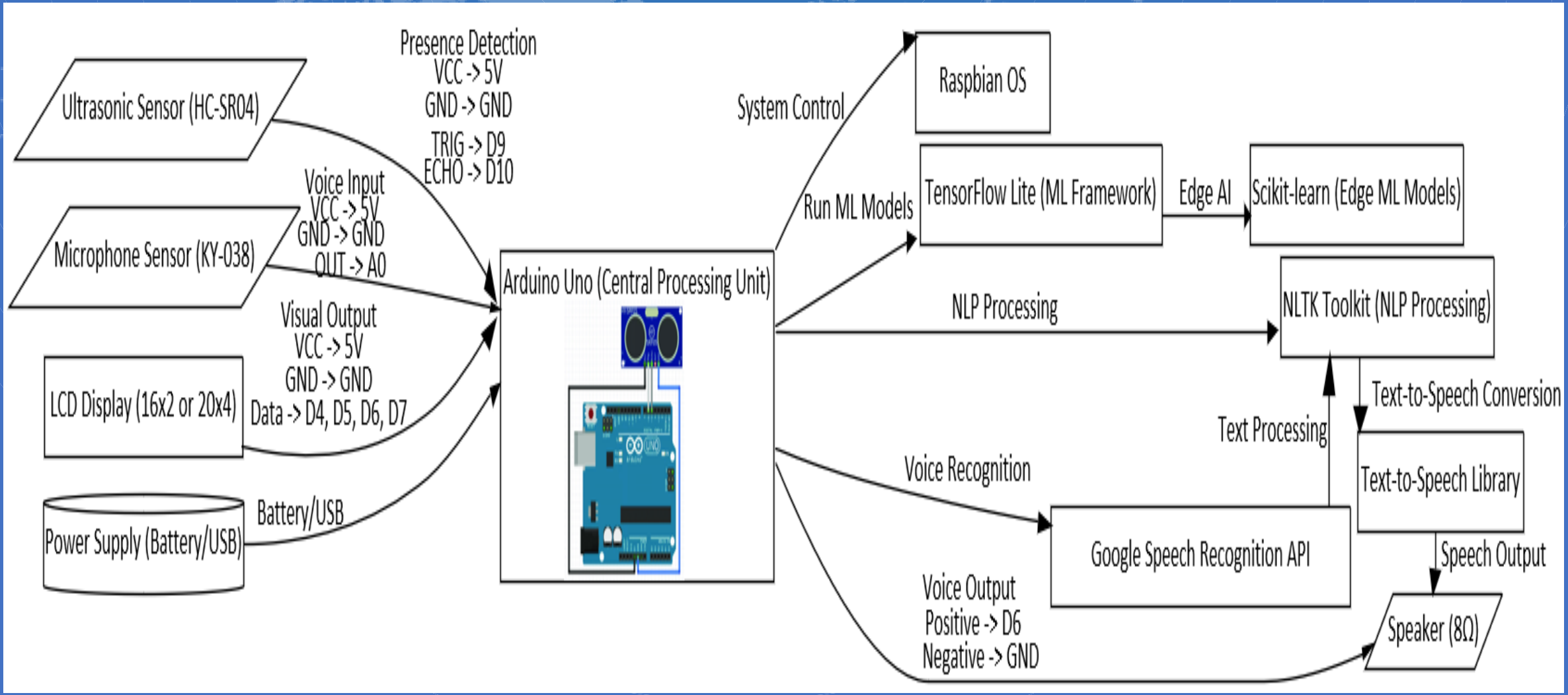
- TensorFlow Lite for running machine learning models on edge devices.
- NLTK toolkit for evaluating and processing human language.
- Google Speech Recognition API for reliable voice-to-text conversion.
- Text-to-speech libraries for natural conversation.

مخطط النظام

- **يُظهر المخطط الخطوات الرئيسية في نظام الإرشاد الذكي:**
- **المدخلات** من خلال المستشعرات فوق الصوتية والميكروفون.
- **معالجة البيانات** باستخدام أطر تعلم الآلة ومعالجة اللغة الطبيعية (NLP)
- **المخرجات** من خلال مكبر الصوت أو مكتبة تحويل النص إلى كلام.



System Architecture



الخاتمة

- يقوم نظام الإرشاد الذكي بتسهيل عملية اختيار التخصص لطلاب الجامعة العربية الأمريكية من خلال الاستفادة من نماذج تعلم الآلة.
- تحسين عملية اتخاذ القرار للطلاب وتعزيز الدقة وقابلية الاستخدام لأنظمة الإرشاد.
- تعزيز كفاءة الإرشاد ومواءمته مع تفضيلات الطلاب.
- تعزيز رضا الطلاب وتحسين عملية اتخاذ القرار.
- تحسين كفاءة الإرشاد للجامعة.
- قابلية التوسع للتطبيقات المستقبلية.

التوصيات

- الاستفادة من خوارزمية Random Forest) لقوتها ودقتها.
- التجربة باستخدام نماذج Deep Learningلاكتشاف العلاقات المعقدة.
- توسيع مجموعة البيانات لتشمل ميزات إضافية (مثل الدخل وفرص العمل).