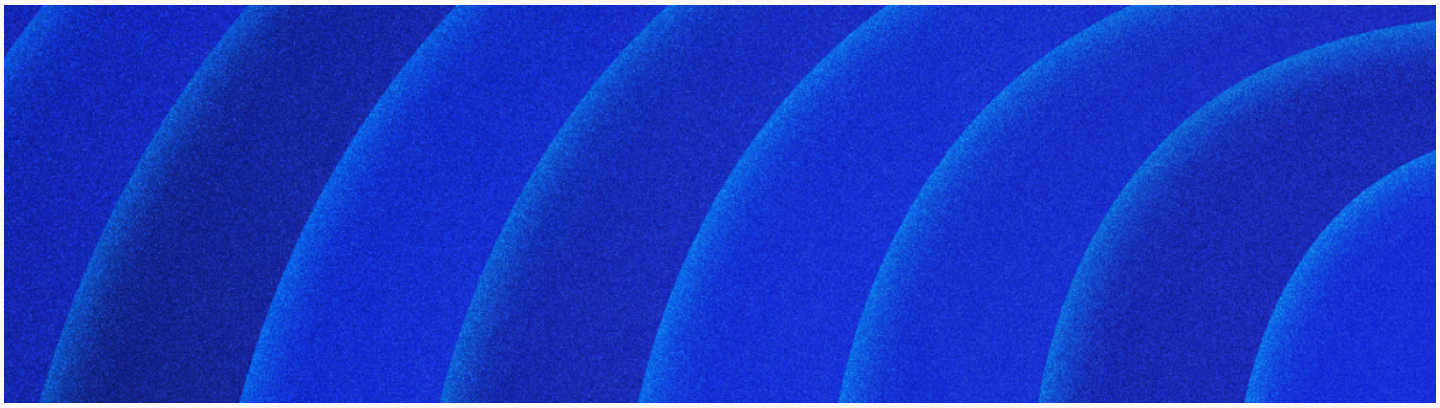

Quintessence AI

Mar 17, 2025

Final Portfolio Project

Presented by Ambareen Baig

An AI-powered
Startup Screening Application

01

EXECUTIVE SUMMARY

Quintessence AI is an AI-powered startup screening tool that automates application evaluation for incubators and investors. It combines **Predictive Modeling** and **LLM-based qualitative analysis** to assess startups based on structured criteria, reducing manual workload and bias. The tool assigns AI-driven selection scores, predicts startup success, and provides insights via an interactive **Streamlit dashboard**. Deployed with **Random Forest for structured data** and **GPT-3.5 for textual analysis**, it enhances decision-making efficiency. By ensuring **data-driven, consistent, and fair evaluations**, Quintessence AI transforms the startup selection process.

02

PROBLEM

- Describe pain points and challenges your target customers face
- Quantify the problem's impact with data, if possible

SOLUTION

- Reintroduce the product or service
- Explain the ways it uniquely solves the problem
- Highlight key differentiators

03

INTRODUCTION

Problem Statement

Startup incubators, accelerators, and investors receive a high volume of applications, making the selection process time-intensive, prone to bias, and inconsistent. The current manual review process relies on subjective evaluation methods, which vary based on the experience and background of individual reviewers. This inconsistency can lead to the selection of businesses that may not have the highest potential for success while overlooking promising startups. Additionally, without predictive analytics, evaluators lack a structured way to anticipate which startups are most likely to succeed based on historical data and relevant business attributes.

04

PROJECT OBJECTIVES

What will this solution achieve?

The goal of this project is to develop **Quintessence AI**, an AI-powered application screening model that automates the evaluation process by leveraging **Large Language Models (LLMs) and Predictive Modeling**. The model is designed to:

- Provide a standardized, data-driven evaluation process that reduces subjectivity and enhances consistency in startup selection.
- Analyze structured data, such as financial performance and market traction, to predict startup success based on historical selection patterns.
- Process unstructured textual data, such as business descriptions and value propositions, to derive insights into innovation potential and industry fit.
- Improve the efficiency of the selection process by reducing the manual workload required for initial application reviews.

05 TECHNICAL APPROACH

Methodology

Quintessence AI consists of two primary components that work together to evaluate startup applications:

- ***Predictive Modeling for Structured Data Analysis***
    - A supervised machine learning model is trained on past startup applications to classify businesses as either likely to be selected or rejected. The predictive model takes into account structured numerical and categorical features such as industry, revenue, funding history, and team expertise.
- ***LLM-Based Analysis for Unstructured Data***
    - A language model processes textual descriptions of business models, market strategies, and value propositions to generate qualitative insights. This allows evaluators to gain deeper context about why a particular startup may or may not align with the selection criteria.

Tools and Technologies

The project was implemented using a combination of machine learning and natural language processing tools, including:

- **Python** for data processing and model development.
- **Scikit-learn** for training and optimizing the predictive model.
- **OpenAI GPT-3.5 API** for analyzing textual data and answering evaluator queries.
- **Pandas and NumPy** for data cleaning and feature engineering.
- **Streamlit** for creating an interactive dashboard that displays AI-generated evaluations.
- **SQLite** for managing structured data storage.

06

DEVELOPMENT PROCESS

Data Consolidation & Preprocessing

Tools and Technologies

- **Historical Startup Application Data:** A dataset containing information on startup applications, including past acceptance and rejection decisions, industry classification, and key business attributes.
- **Business Performance Data:** A dataset tracking post-incubation business outcomes, including revenue growth, market expansion, and investment secured.

Data Cleaning and Feature Engineering

- Business names were standardized to ensure consistency across datasets, preventing duplicate records from being misclassified.
- Unique business identifiers were assigned to allow seamless cross-referencing between datasets.
- Missing values were handled through imputation techniques to ensure data completeness.
- Additional features were engineered, including a **Selection Likelihood Score**, which quantifies how closely an applicant aligns with past successful selections.

Model Training and Implementation

Predictive Model (Random Forest Classifier)

- The model was trained on structured features, such as market potential, team expertise, and financial indicators, to classify businesses as either accepted or rejected.
- The model outputs a confidence score, indicating the likelihood that a startup would be selected based on past application patterns.

AI Assistant (GPT-3.5-Based Insights Generator)

- The assistant analyzes qualitative data, extracting themes from business descriptions and industry trends.

- Evaluators can interact with the AI assistant to ask questions about selection patterns, common success factors, and reasons behind past decisions.

Model Deployment and User Dashboard

- The AI system is deployed using **Streamlit**, providing an intuitive interface where users can enter startup details and receive AI-generated selection scores.
- The dashboard includes:
    - An application screening module that assigns an AI-driven score based on structured criteria.
    - A ranking system that prioritizes applications based on predicted selection likelihood.
    - A data insights module that visualizes trends in past selections and industry success rates.
    - An AI assistant that allows users to query historical selection patterns and industry insights.

🚀 Quintessence AI: Business Selection Tool

✍️ **Manual Entry (if no CSV uploaded)**

📌 Value Proposition

Select Business Sector

Healthcare_and_Allied_services

📌 Market & Growth Potential

Annual Revenue

10000

Number of Employees

5

Investment Received

1000

Number of Clients

100

Number of Rural Producers Supported

0

📌 Team & Expertise

Highest Level of Education

PhD

Age of Founder

35

Gender of Founder

Male

Predict Selection Status

📝 AI-Based Analysis of Business Descriptions

💬 Ask the AI Assistant

07

OUTCOMES & RESULTS

What can you achieve with this Quintessence AI?

Key Achievements

- The AI-powered screening model significantly reduced the time required for initial application reviews by automating the evaluation process.
- The predictive model achieved an **accuracy of 57%**, closely aligning with past human evaluator decisions.
- The LLM-based assistant provided valuable contextual insights, improving transparency in selection decisions.
- The system enabled evaluators to identify high-potential startups more efficiently, improving consistency in decision-making.

Performance Evaluation

57% Model Accuracy

69% Precision (Rejected Cases)

68% Recall (Rejected Cases)

34% F1 Score

0% Precision, Recall, and F1-score for Accepted Cases: The model did not correctly classify any accepted businesses

08 CHALLENGES & SOLUTIONS

Challenges Faced

- ***Processing*** **text-heavy applications**: The initial model struggled to analyze unstructured business descriptions effectively.
- ***Balancing*** **AI predictions with human judgment**: AI-generated scores occasionally differed from human evaluations.
- ***Overfitting*** **in the predictive model**: The model initially classified too many startups as likely to be accepted.
- ***Bias*** **in training data**: Past selection decisions had inherent biases that the model initially replicated.

Solutions Implemented

- **Incorporated LLM processing** to extract key insights from textual data, improving qualitative analysis.
- **Implemented threshold adjustments** to align AI-generated selections more closely with historical human decisions.
- **Applied cross-validation techniques** to prevent overfitting and ensure the model generalized well to new applications.
- **Introduced bias correction mechanisms** to mitigate historical biases in the training data.

10 CONCLUSION

Quintessence AI successfully automates startup application screening by combining predictive analytics with AI-powered qualitative insights. The system enhances efficiency, improves fairness, and provides a data-driven approach to startup evaluation. By reducing subjectivity in the selection process and enabling evaluators to make more informed decisions, this tool represents a significant advancement in AI-driven business intelligence for startup ecosystems.

11 GITHUB

Github Repository Link: Click Here

Live Streamlit App Link: Launch Quintessence AI