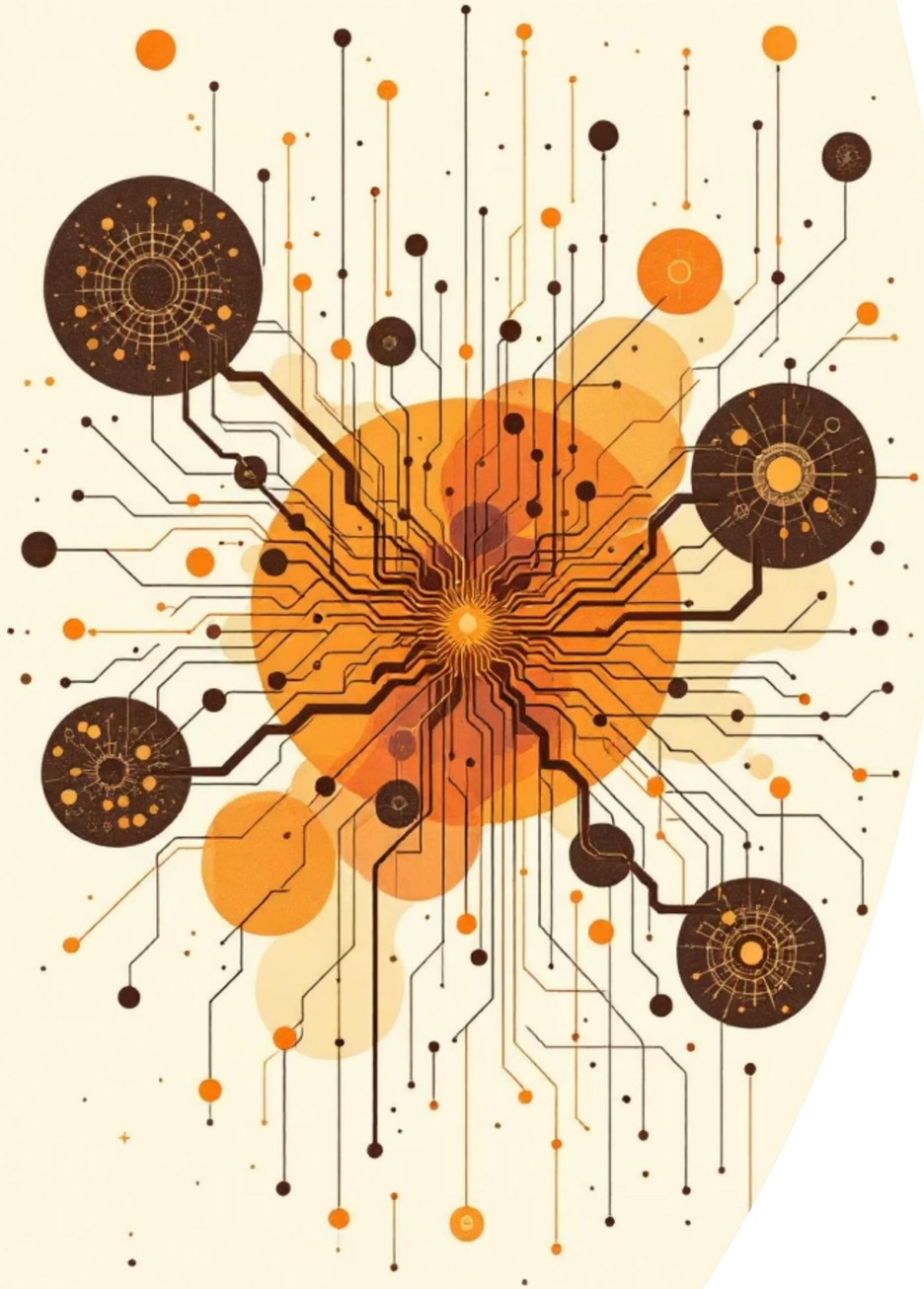


Udyogaa AI — Hiring Intelligence Prototype

AI-powered backout prediction and hiring intelligence system designed to transform recruitment from reactive to predictive.

Aditya Sawarkar | Data Scientist

Hiring Challenges in Modern Recruitment

Traditional recruitment processes struggle with unpredictability and inefficiency in today's competitive talent market.

Offer Drop-Offs

The risk of candidates accepting offers but failing to join creates significant uncertainty in hiring timelines and budget planning. Organizations lose time, resources, and momentum when offers fall through.

Unpredictable Joining

Manual processes and human judgment cannot reliably forecast which candidates will actually commit to joining the organization, leading to repeated setbacks.

Screening Inefficiency

Recruiting teams waste countless hours manually reviewing resumes and LinkedIn profiles, with inconsistent evaluation criteria that slow down the hiring process.

Data Gap

Hiring workflows lack predictive analytics to guide decisions, forcing teams to rely on intuition rather than objective, data-driven insights.

AI-Powered Hiring Intelligence System

1

Raw Data

Candidate profiles, historical hiring data, and communication patterns fed into the system

2

Feature Engineering

Extracting meaningful signals: experience duration, communication responsiveness, technical scores

3

Multi-Model Training

Testing multiple algorithms to identify highest-performing predictive models

4

Best Model Selection

Automatic selection based on performance metrics, deployed for production use

This integrated system combines predictive analytics and semantic intelligence to deliver actionable hiring insights.

Backout Risk Prediction Model

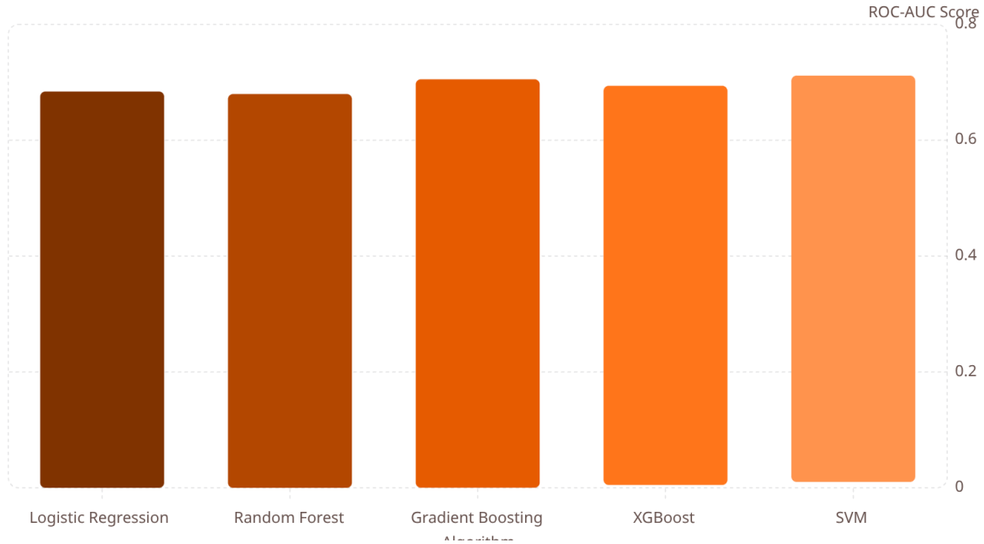
How It Works

A binary classification model distinguishes between candidates who will join (1) versus those who will back out (0). The system analyzes hundreds of data points to calculate risk probability.

Model Selection Strategy

Five classification algorithms trained and validated on historical hiring data. Best model selected automatically using ROC-AUC metric, which evaluates how well each model distinguishes between joining and non-joining candidates.

Performance Comparison



Selected Model: Gradient Boosting achieved the highest ROC-AUC score of 0.7034, showing strongest predictive capability for backout risk.

Business-Oriented Candidate Scoring

Weighted Scoring Framework

Candidates receive a comprehensive score blending technical capability, communication quality, risk assessment, and stability factors.

$$\textit{Final Score} = 0.4 \times \textit{Technical} + 0.2 \times \textit{Communication} + 0.2 \times (1 - \textit{Risk}) + 0.2 \times \textit{Stability}$$

Technical Skills

40% weight

Assessment scores, certifications, experience relevance

Communication

20% weight

Responsiveness, clarity, and engagement quality

Risk Inverse

20% weight

(1 - Backout Probability) rewards low-risk candidates

Stability

20% weight

Job tenure, career trajectory consistency

The score ranges from 0 to 100, providing recruiters an intuitive metric to prioritize candidates.

Semantic Job-Resume Matching

Context-Aware Matching

Traditional keyword matching fails to capture meaning. Our system uses Sentence Transformers to generate semantic embeddings that understand context and relationships between words.

The process converts both job descriptions and resumes into high-dimensional vectors. Cosine similarity then measures how closely the candidate's profile matches the role's requirements.

Matching Score Range: 0 to 1, where higher values indicate stronger alignment between candidate qualifications and job requirements.

01

Job Description

Requirement posted or internal position defined

02

Embedding Generation

Sentence Transformers encode text as vectors

03

Cosine Similarity

Calculate alignment between job and candidate vectors

04

Matching Score

Similarity measure ranked 0 to 1

Streamlit Dashboard Interface

Risk Percentage

0–100% probability visualization

Risk Category

Low / Medium / High classification

Final Score

0–100 composite assessment

Joining Probability

Machine learning predicted probability

Business Recommendation

Actionable next step guidance

Dashboard Features

- Interactive visualization of risk factors
- Color-coded urgency indicators
- Confidence intervals for predictions
- Historical performance comparisons
- Exportable reports for stakeholder review

Business Value Delivered

Transform recruitment from reactive crisis management to proactive, data-driven strategic planning.

1

Reduce Offer Drop-Offs

Flag high-risk candidates before extending offers, minimizing backout-related costs and reputational damage

2

Improve Hiring Predictability

Forecast joining probabilities with measurable confidence, enabling reliable headcount and budget planning

3

Increase Recruiter Efficiency

Smart recommendations allow recruiters to focus on high-potential candidates

4

Automate Resume Screening

Semantic matching and scoring eliminates manual document review, processing hundreds of applications in seconds

5

Enable Data-Driven Decisions

Replace gut feelings with objective metrics, building a culture of analytical recruitment

> From Reactive Hiring → Predictive Hiring

Key Engineering Design Principles

Architecture Built for Production

Multi-Model Comparison Pipeline

Grid search across algorithms with standardized evaluation metrics. Enables performance comparisons.

Dynamic Model Deployment

Deploy new model versions without downtime as performance improves

Modular Project Structure

Separation of data loading, feature engineering, model training, and scoring logic

Integrated ML + NLP

Unified codebase for predictive modeling and semantic matching

Business Interpretation Layer

Transforms raw probabilities into actionable risk categories and recommendations

Future Enhancements Roadmap

Advanced Validation

Hyperparameter optimization and k-fold cross-validation.

Transparency Layer

SHAP values explaining model decision factors

Adaptive Thresholds

Organization-specific risk scoring adjustments

Ranking Interface

Compare multiple candidates across offers simultaneously

HR System Integration

REST API endpoints for Workday, Greenhouse, and Lever