

# Startup-Investor Connect using ML

**Mrs. P. Niharika<sup>1</sup>, Sreehasa Pendyala<sup>2</sup>, Sri Keerthi Bavanila<sup>3</sup>,  
Sarayu Shiramshetti<sup>4</sup>, Pranay Kumar Boindla<sup>5</sup>**

Assistant Professor, Department of CSE (Data Science), ACE Engineering College, India<sup>1</sup>  
IV B.Tech, Department of CSE (Data Science), ACE Engineering College, India<sup>2,3,4,5</sup>  
niharika4jan@gmail.com<sup>1</sup>, pendyalasreehasa@gmail.com<sup>2</sup>, bavanilasrikeerthi@gmail.com<sup>3</sup>,  
sarayushiramshetti22@gmail.com<sup>4</sup>, pranayboindla@gmail.com<sup>5</sup>

**Abstract:** *The startup ecosystem has been growing rapidly over the last few years. This has led to an increasing demand for efficient mechanisms for matching startup companies with suitable investors. Current discovery processes including attending networking events and manually screening and filtering basic profile attributes do not successfully identify optimal matches because they conduct insufficient analysis of compatibility-related elements. Our research proposes Startup-Investor Connect that is a machine learning recommendation system to facilitate the startup-investor matching process. The system collects structured profile data including sector, funding stage, funding requirements, investment range, team size and location. Data preprocessing and feature engineering techniques are applied to convert raw profile data into compatibility metrics which contain sector match and funding stage match and investment range alignment. Our team experimented with various machine learning models including Random Forest, Logistic Regression, K-Nearest Neighbors and Naive Bayes to determine the most accurate prediction model. The Random Forest classifier performed the most robustly and was selected for compatibility prediction. The recommendation engine orders startup-investor pairs based on predicted compatibility scores and shows the top matches through a web interface. Our proposed system offers a scalable, data-driven solution for enhancing investor discovery and facilitating more efficient startup funding ecosystems*

**Keywords:** Machine Learning (ML), Recommendation Systems, Random Forest Classifier, Compatibility Scoring, Startup Ecosystem, Venture Capital Analytics, Startup-Investor Matching

## I. INTRODUCTION

Startups are the engines of innovation, economic growth and technological progress. While startups face the challenge of finding investors whose business needs, funding and growth stage match, investors cope with the task of finding startups whose investment method and risk tolerance match.

Startups traditionally go about finding investors using networking events, incubator programs and personal referrals. These methods help create initial exposure to the target audiences, but require manual work to achieve their full potential. While online investment platforms are able to filter results based on industry and funding stage, they do not offer any intelligent analysis of the compatibility between the startup's needs and investor's preferences.

Recent developments in machine learning and recommendation systems enable us to enhance this process using data-driven approaches for matchmaking. Machine learning models analyze structured profile data to identify patterns that indicate potential matches between startups and investors.

This paper presents Startup-Investor Connect that proposes a machine learning-based recommendation system that works by evaluating the startup and investor profiles to compute compatibility scores and ranked suggestions. The system utilizes data preprocessing and feature engineering along with machine learning prediction and recommendation ranking for making meaningful connections that are beneficial to both investors and startups.



### **Contributions**

Our contributions are summarized as follows:

- We have designed a recommendation system for matchmaking between startup and investors through machine learning.
- We have implemented various compatibility features and compared several machine learning models (Random Forest, Logistic Regression, K-Nearest Neighbors, and Naive Bayes) to find the most accurate prediction algorithm.
- We have implemented a web application that lists recommendations for both startup and investors in order of relevance.

## **II. EXISTING SYSTEM AND DRAWBACKS**

### **A. Existing System**

Current startup-investor discovery platforms primarily rely on basic search and networking mechanisms.

1. Static profile platforms permit startup and investor users to conduct profile searches by applying basic filters which include industry and funding stage and location.
2. Networking-driven discovery occurs through startup events demo days accelerator programs and personal referrals.
3. Keyword or tag-based matching systems use industry tags to classify startups while offering users restricted options for filtering results.

### **B. Drawbacks**

These approaches suffer from several limitations:

Lack of personalized recommendations

High dependence on manual searching

Inefficient matching when datasets grow large

No compatibility scoring mechanism

Limited analysis of investor preferences and startup characteristics.

As a result, startups may struggle to identify suitable investors, while investors must manually evaluate large numbers of potential opportunities.

## **III. PROPOSED SYSTEM AND ADVANTAGES**

### **A. Proposed System**

The recommended system is a machine learning recommendation system that allows a startup to be matched with an appropriate investor through compatibility matching. The system obtains at least a structured profile of a startup and a structured profile of an investor, where the structured profile includes information about a sector, a funding stage, a funding requirement, an investment range, a team size, and a location. The system determines a feature engineering of the acquired data and generates a set of compatibility indicators that represent the compatibility between the startup needs and the investor needs.

The machine learning model evaluates the compatibility indicators and determines a set of the most compatible startup-investor pairs. The recommendation engine ranks all possible matches based on the calculated compatibility scores and shows the ranked matches to a user through a web interface. A startup can use the platform to find investors who meet its funding needs. An investor can use the platform to find startups that match its investment preferences.

### **B. Advantages**

The proposed system offers several benefits:

Personalized startup-investor matching

Data-driven compatibility analysis

Reduced manual effort in investor discovery



Faster and more efficient matchmaking  
Scalable architecture capable of handling large datasets.

#### **IV. SYSTEM MODULES**

##### **1. User Registration and Profile Creation**

Startup and Investor can register their accounts and create their user profiles. Startup profiles contain information about the startup's sector of business and current funding and required funding and number of team members. Investor profiles contain information about their preferred sector of investment, investment range and risk appetite.

##### **2. Data Processing and Feature Engineering**

Team has performed the necessary data cleaning and data normalization to generate the machine learning features from the raw data. The system generates necessary compatibility metrics which includes sector matching and funding matching and stage matching and team matching.

##### **3. Machine Learning Module for Prediction**

The module uses the Random Forest model which has been trained to predict how well a startup profile matches with an investor profile.

Random Forest was chosen as the final prediction model. Random Forest can handle categorical and numerical features and can produce a more stable prediction by averaging the projections of multiple decision trees.

##### **4. Recommendation Engine**

The recommendation engine generates its best matches for a given pair of startup and investor by matching the compatibility scores for each pair.

##### **5. Result Visualization**

The results are shown to the user via the web interface where user can view matched startups or investors along with their compatibility scores.

#### **V. SYSTEM ARCHITECTURE**

The system follows a layered architecture consisting of three layers:

1. Presentation Layer: Web interface where users register, create profiles, and view recommendations.
2. Application Layer: Implements the core logic using the Flask framework, including preprocessing, prediction, and recommendation generation.
3. Data Layer: Stores startup datasets, investor datasets, user profiles, and the trained machine learning model file.



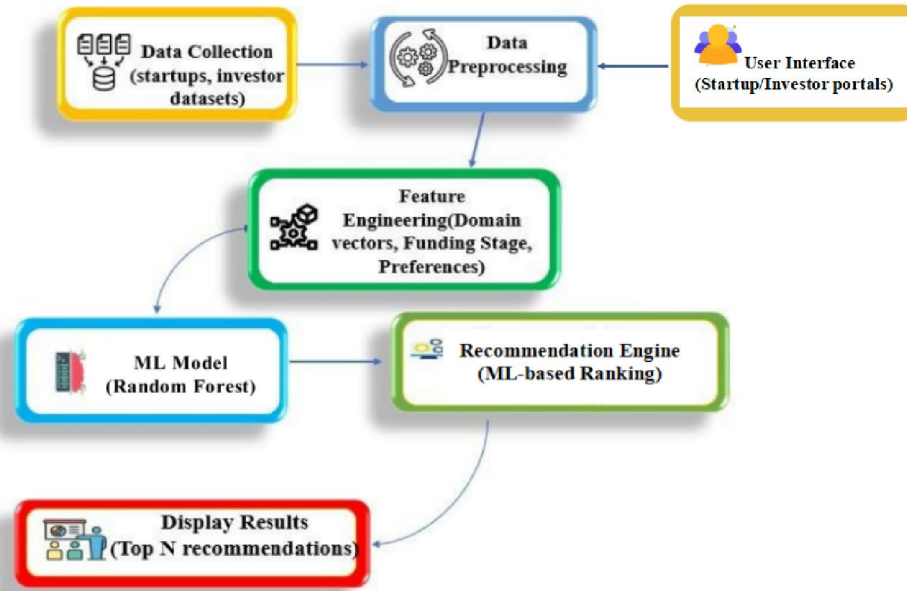


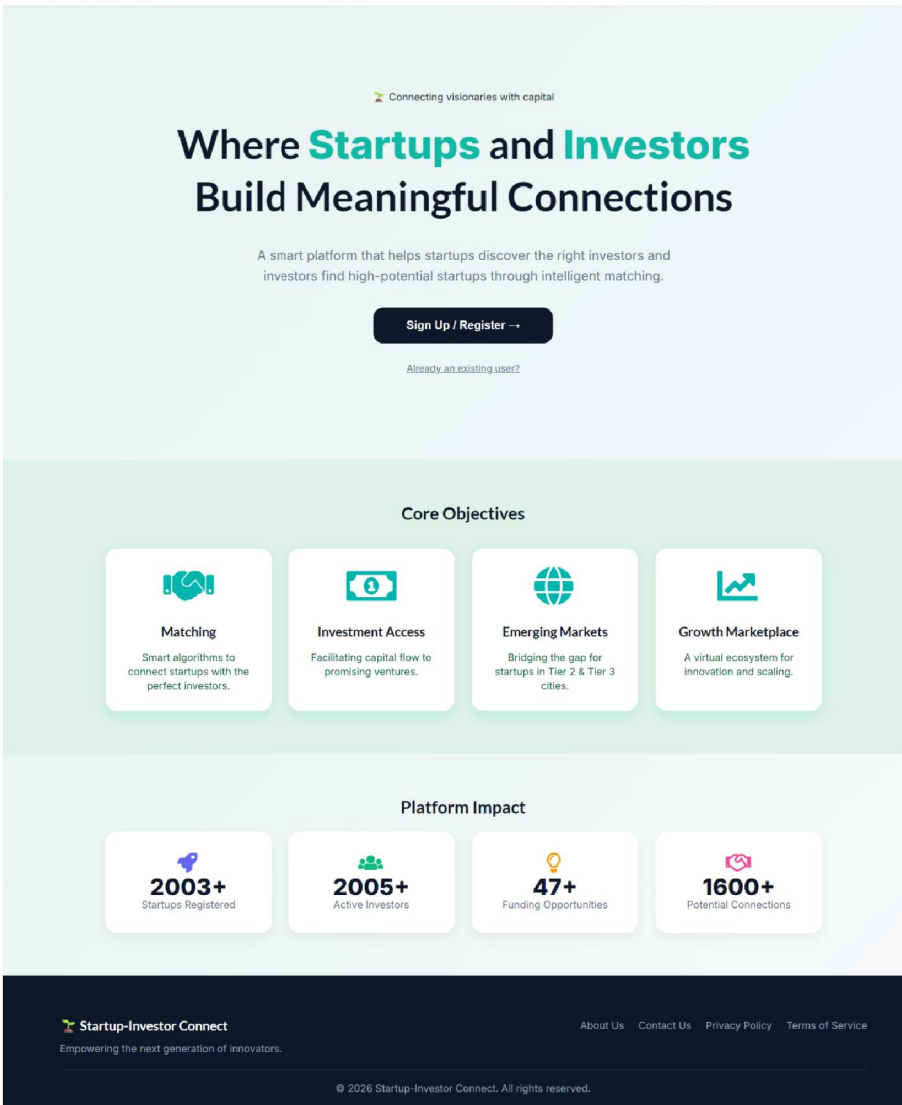
Fig-1: System Architecture

## VI. RESULTS AND OUTPUTS

The system provides an interactive web platform where users can access recommendation results. Major outputs include:

- startup registration interface
- investor registration interface
- recommended investors for each startup
- recommended startups for each investor
- compatibility scores for each match.





**Startup-Investor Connect**

Connecting visionaries with capital

## Where **Startups** and **Investors** Build Meaningful Connections

A smart platform that helps startups discover the right investors and investors find high-potential startups through intelligent matching.

[Sign Up / Register →](#)

[Already an existing user?](#)

### Core Objectives

- Matching**  
Smart algorithms to connect startups with the perfect investors.
- Investment Access**  
Facilitating capital flow to promising ventures.
- Emerging Markets**  
Bridging the gap for startups in Tier 2 & Tier 3 cities.
- Growth Marketplace**  
A virtual ecosystem for innovation and scaling.

### Platform Impact

- 2003+** Startups Registered
- 2005+** Active Investors
- 47+** Funding Opportunities
- 1600+** Potential Connections

**Startup-Investor Connect**  
Empowering the next generation of innovators.

[About Us](#) [Contact Us](#) [Privacy Policy](#) [Terms of Service](#)

© 2026 Startup-Investor Connect. All rights reserved.

Fig-2: Interface for Registration / Login



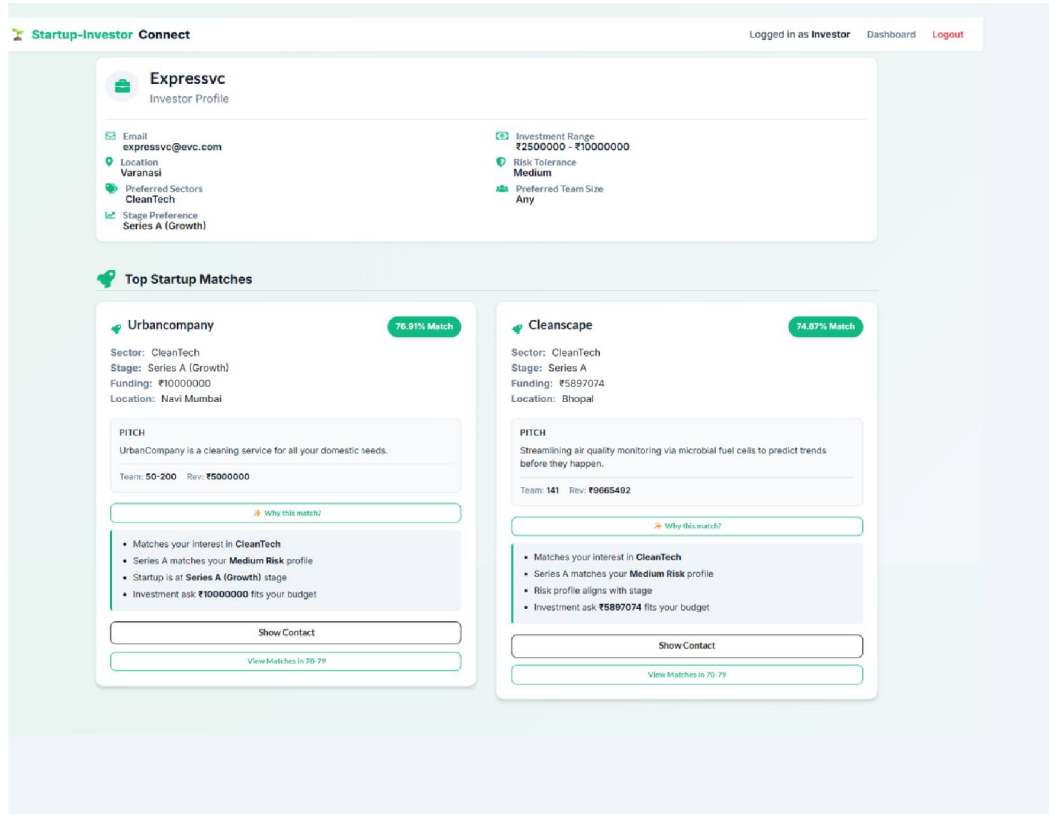
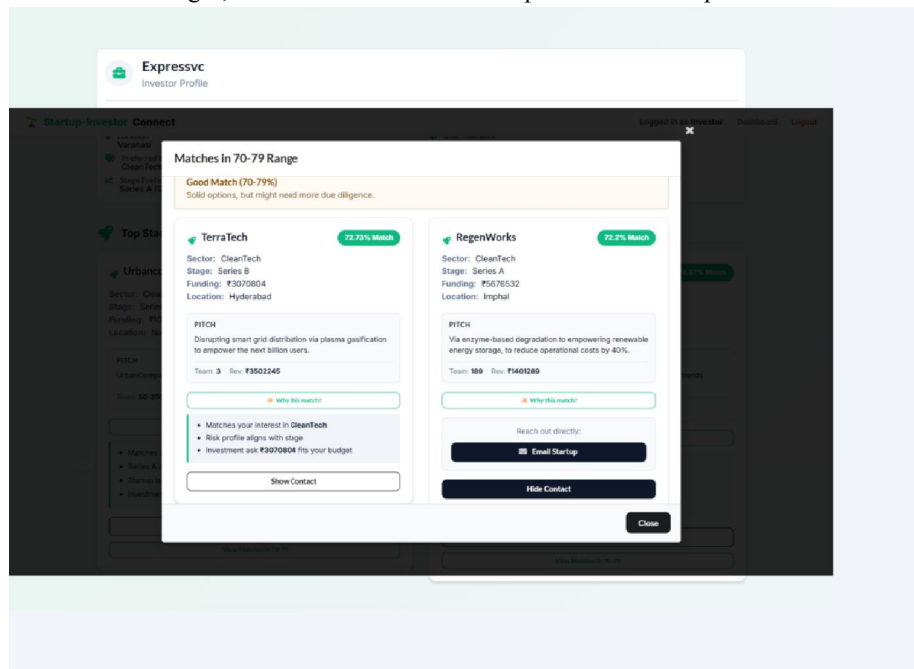
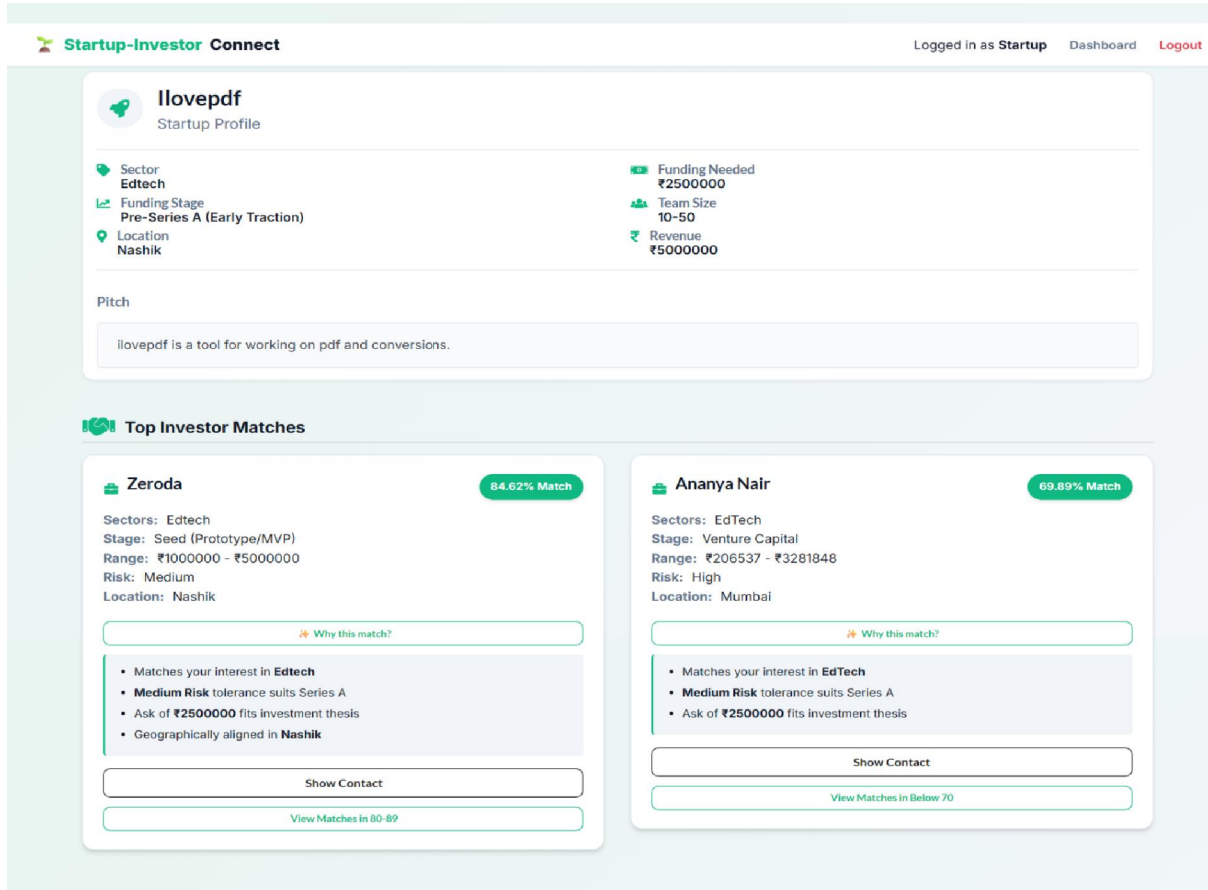


Fig-3,4: Recommendation of Startups in the Investor portal





**Startup-Investor Connect** | Logged in as **Startup** | Dashboard | Logout

### Ilovepdf

Startup Profile

- Sector: Edtech
- Funding Stage: Pre-Series A (Early Traction)
- Location: Nashik
- Funding Needed: ₹2500000
- Team Size: 10-50
- Revenue: ₹5000000

Pitch: Ilovepdf is a tool for working on pdf and conversions.

### Top Investor Matches

**Zeroxa** 84.62% Match

Sectors: Edtech  
Stage: Seed (Prototype/MVP)  
Range: ₹1000000 - ₹5000000  
Risk: Medium  
Location: Nashik

Why this match?

- Matches your interest in **Edtech**
- Medium Risk** tolerance suits Series A
- Ask of **₹2500000** fits investment thesis
- Geographically aligned in **Nashik**

Show Contact

View Matches in 80-89

**Ananya Nair** 69.89% Match

Sectors: EdTech  
Stage: Venture Capital  
Range: ₹206537 - ₹3281948  
Risk: High  
Location: Mumbai

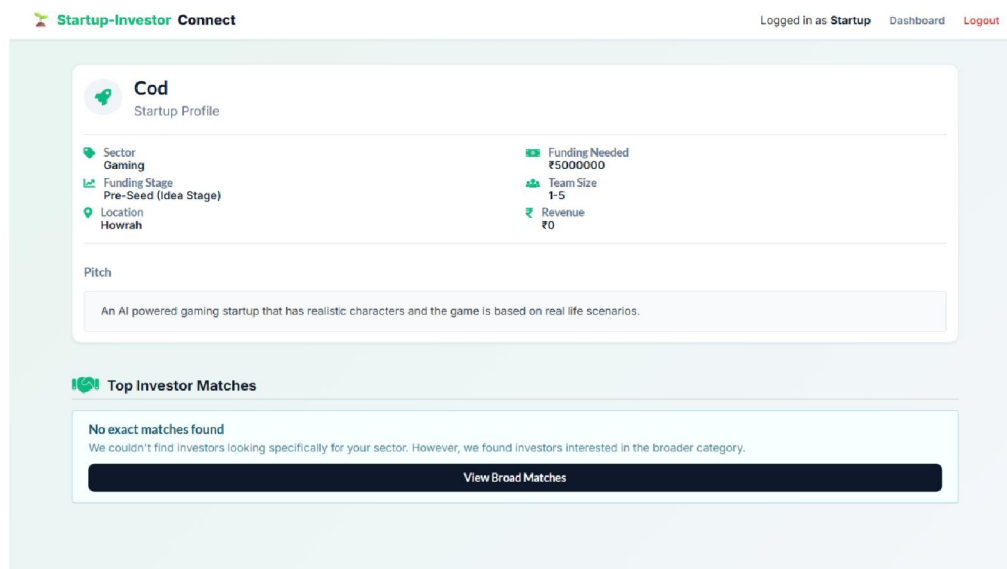
Why this match?

- Matches your interest in **EdTech**
- Medium Risk** tolerance suits Series A
- Ask of **₹2500000** fits investment thesis

Show Contact

View Matches in Below 70

Fig-5: Recommendation of Investors in the Startup portal (Interface when matches exist)



**Startup-Investor Connect** | Logged in as **Startup** | Dashboard | Logout

### Cod

Startup Profile

- Sector: Gaming
- Funding Stage: Pre-Seed (Idea Stage)
- Location: Howrah
- Funding Needed: ₹5000000
- Team Size: 1-5
- Revenue: ₹0

Pitch: An AI powered gaming startup that has realistic characters and the game is based on real life scenarios.

### Top Investor Matches


No exact matches found

We couldn't find investors looking specifically for your sector. However, we found investors interested in the broader category.

View Broad Matches

Fig-6: Recommendation of Investors in the Startup portal (Interface when no matches exist)





**Cod**  
Startup Profile

---

**Sector**  
Gaming

**Funding Stage**  
Pre-Seed (Idea Stage)

**Location**  
Howrah

**Funding Needed**  
₹5000000

**Team Size**  
1-5


**Revenue**  
₹0

---

**Pitch**

An AI powered gaming startup that has realistic characters and the game is based on real life scenarios.

---



**Top Investor Matches**

**No exact matches found**  
We couldn't find investors looking specifically for your sector. However, we found investors interested in the broader category.

**Matches in Broad Sector**

**Kavya Iyer**  
Delhi

**Investment Range:** ₹172988 - ₹3688429

**Sectors:** SaaS

59.64% Match

---

**Analysis:**  
Matches broad interest in **SaaS** • **High Risk** tolerance matches early stage • Capital requirement is close to target range

Why this match?

- Matches broad interest in SaaS
- High Risk tolerance matches early stage
- Capital requirement is close to target range

Contact Investor

---

**Ananya Gupta**  
Hyderabad

**Investment Range:** ₹88888 - ₹3338089

**Sectors:** SaaS

**Email:**  
ananyagupta1852@gmail.com

---

**Analysis:**  
Matches broad interest in **SaaS** • **High Risk** tolerance matches early stage • Capital requirement is close to target range

Why this match?

- Matches broad interest in SaaS
- High Risk tolerance matches early stage
- Capital requirement is close to target range

Hide Contact

Fig-7: Recommendation of Investors in the Startup portal (Interface for broader matches)



## **VII. CONCLUSION AND FUTURE SCOPE**

### **A. Conclusion**

We propose Startup-Investor Connect, a machine learning recommendation system to improve the matching between investors and startups. Structured profile information is fed into feature engineering techniques and machine learning models to calculate compatibility scores for matching the startups with the investors.

We evaluated several machine learning models and chose Random Forest as our model of preference because it produces stable results and works with both types of data (categorical and numerical). The recommendation engine employs its ranking system to assess potential matches and present them via an online interface.

The proposed method shows how machine learning can help in discovering investors, reduce manual efforts and help foster startup funding ecosystems.

### **B. Future scope:**

Future improvements for the system include:

incorporating natural language processing for analyzing startup pitch descriptions

using graph-based recommendation techniques

deploying the system as a scalable cloud-based platform.

## **REFERENCES**

- [1] Xu, S., Lei, Z., & Yang, F., "Recommending Investors for New Startups using Graph-Based Learning," 2020. Link: <https://arxiv.org/abs/1912.02962>
- [2] Luef, G., Granitzer, M., & Schedl, M., "A Recommender System for Early-Stage Venture Investment," 2020. Link: <https://doi.org/10.1145/3341105.3375767>
- [3] Xu, J., & Chen, L., "Heterogeneous Network-Based Venture Capital Recommendation Systems," 2023. Link: <https://doi.org/10.3390/math1016289592>
- [4] Sarwar, B., Karypis, G., Konstan, J., & Riedl, J., "Item-Based Collaborative Filtering Recommendation Algorithms," 2001. Link: <https://dl.acm.org/doi/10.1145/371920.372071>
- [5] Aggarwal, C., Recommender Systems: The Textbook, Springer, 2016. Link: <https://link.springer.com/book/10.1007/978-3-319-29659-3>
- [6] Zhao, Z., Zhang, W., & Zhang, F., "Risk-Hedged Venture Capital Recommendation," 2015. Link: <https://doi.org/10.1145/2792838.2800181>
- [7] Razaghzadeh Bidgoli, M., et al., "Startup Success Prediction Using Machine Learning Techniques," 2024. Link: <https://link.springer.com/article/10.1186/s13731-024-00436-x>
- [8] Asemi, A., et al., "An Intelligent Investment Recommendation System using ANFIS and Multimodal ML," 2024. Link: <https://link.springer.com/article/10.1186/s40537-024-00965-y>

