BATCH - 35

PHASE II - FINAL REVIEW

Our Survey Paper Title:

Smart Startup analyzer for Venture Capitals Using Machine Learning

DOI: 10.55041/IJSREM17996

Project Title:

Kickstarter.ai -> A Tool for Investors / Venture Capitalists to assess start-ups

Under The Guidance of: **Prof. Anupama Girish**, Assistant Professor,
CSE dept., DSCE

Dayananda Sagar Institutions
Department of Computer Science Engineering, DSCE

Team Members & Their Respective Contributions:

1. **Mantej Singh Tuli (1DS19CS086) :**
   - Conducted data preprocessing to prepare the data for the model
   - Deployed the machine-learning model
   - Implemented API calls to interact with the deployed model

2. **Sree Chand R (1DS19CS164) :**
   - Contributed towards the front-end development of the website
   - Designed and implemented the user interface
   - Collaborated with the team to incorporate necessary features and functionalities

3. **Shreyas G (1DS19CS202) :**
   - Developed the machine learning model
   - Performed data preprocessing, training, and testing of various ML techniques
   - Collaborated with the team to define clear evaluation metrics and objectives, aligning the model's outputs with the project's requirements.

4. **Dheemanth A N (1DS19CS710) :**
   - Managed the back-end development of the website.
   - Integrated the ML model with the website's infrastructure.
   - Implemented the necessary APIs and endpoints to communicate with the ML model.

ABOUT OUR PROJECT

• The project revolves around the development of an innovative machine learning (ML) model that holds the potential to predict the success or failure of startups.

•By harnessing the power of artificial intelligence (AI), this project aims to revolutionize the venture capital industry, where investors allocate funds to early-stage companies with promising growth prospects.

•By leveraging AI capabilities, the model equips venture capitalists with a powerful tool to make informed investment decisions.

•It empowers them to accurately assess market potential, evaluate risks meticulously, determine fair valuations, and negotiate advantageous terms.

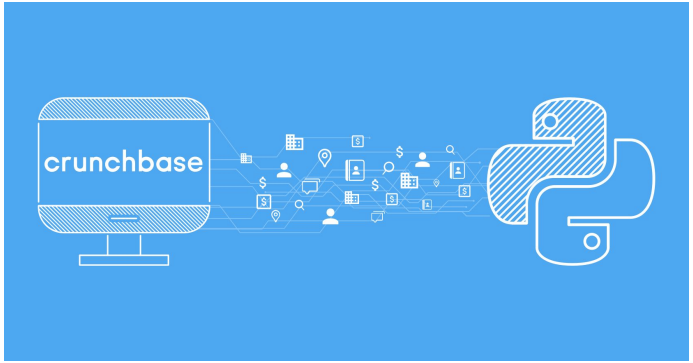
•The ultimate goal of this project is to enhance the efficiency and profitability of venture capital investments.

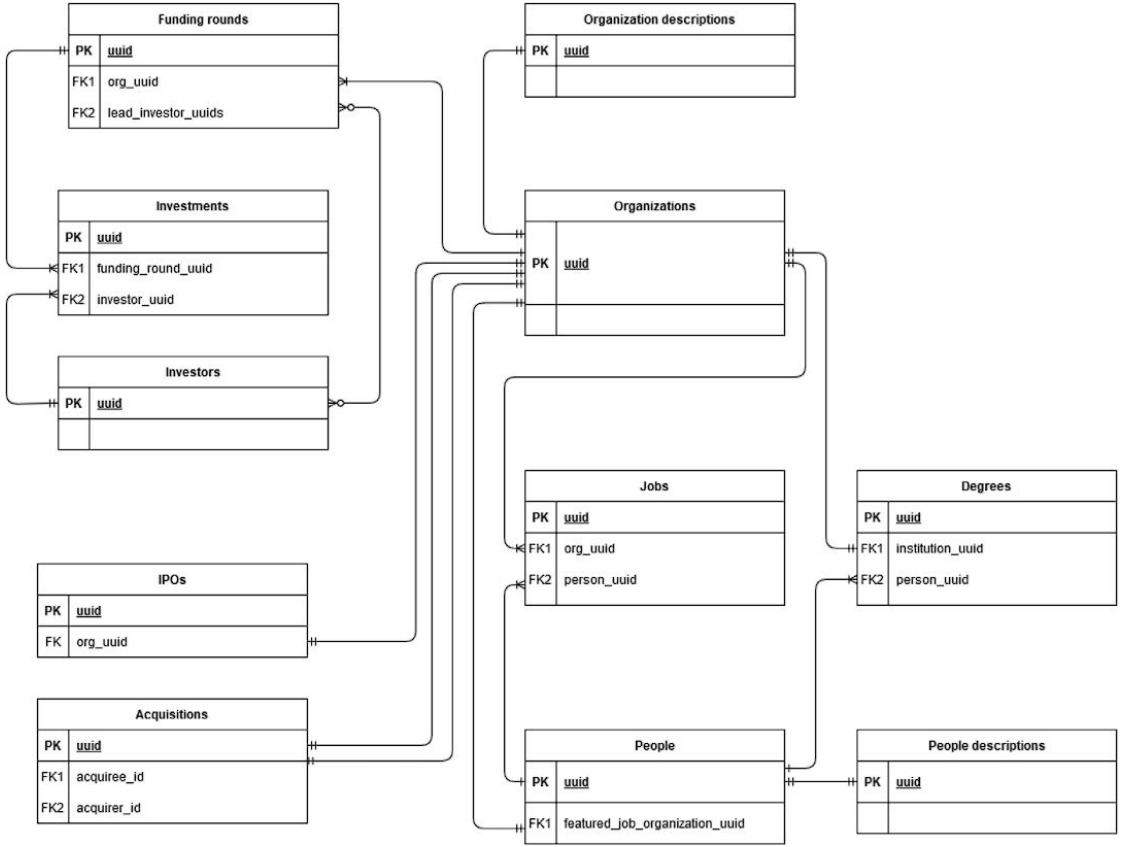
•Its outcome is a sophisticated ML model that empowers venture capitalists to make well-informed decisions, ultimately shaping the success stories of startups and revolutionizing the future of the venture capital industry. It is designed to provide data-driven predictions on whether a startup will achieve success, indicated by factors such as an IPO transition or acquisition by a larger company, or experience failure.

Our Test Product is called - “KICKSTARTER.AI”

DATA

Crunchbase, a vast repository of startup data, houses information on millions of companies, investors, and funding rounds. With a database spanning over 1.5 million organizations and 12 million funding rounds, Crunchbase provides a comprehensive view of the entrepreneurial ecosystem. Its expansive dataset, constantly updated with new information, equips our project with a vast pool of insights to analyze and extract meaningful patterns. By tapping into this wealth of numerical data, we leverage Crunchbase's extensive coverage to enhance the accuracy and reliability of our predictive model.



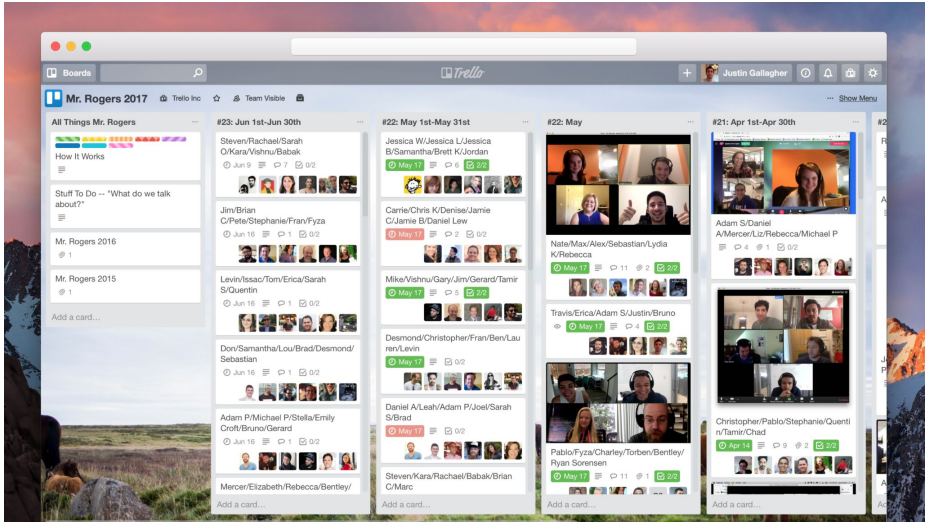

Funding rounds
PK uuid
FK1 org_uuid
FK2 lead_investor_uuids
Organization descriptions
PK uuid
Investments
PK uuid
FK1 funding_round_uuid
FK2 investor_uuid
Organizations
PK uuid
Investors
PK uuid
Jobs
PK uuid
FK1 org_uuid
FK2 person_uuid
Degrees
PK uuid
FK1 institution_uuid
FK2 person_uuid
IPOs
PK uuid
FK org_uuid
Acquisitions
PK uuid
FK1 acquiree_id
FK2 acquirer_id
People
PK uuid
FK1 featured_job_organization_uuid
People descriptions
PK uuid

PROGRESS TRACKING OF OUR PROJECT VIA A PROJECT MANAGEMENT TOOL

ATLASSIAN

+

Trello

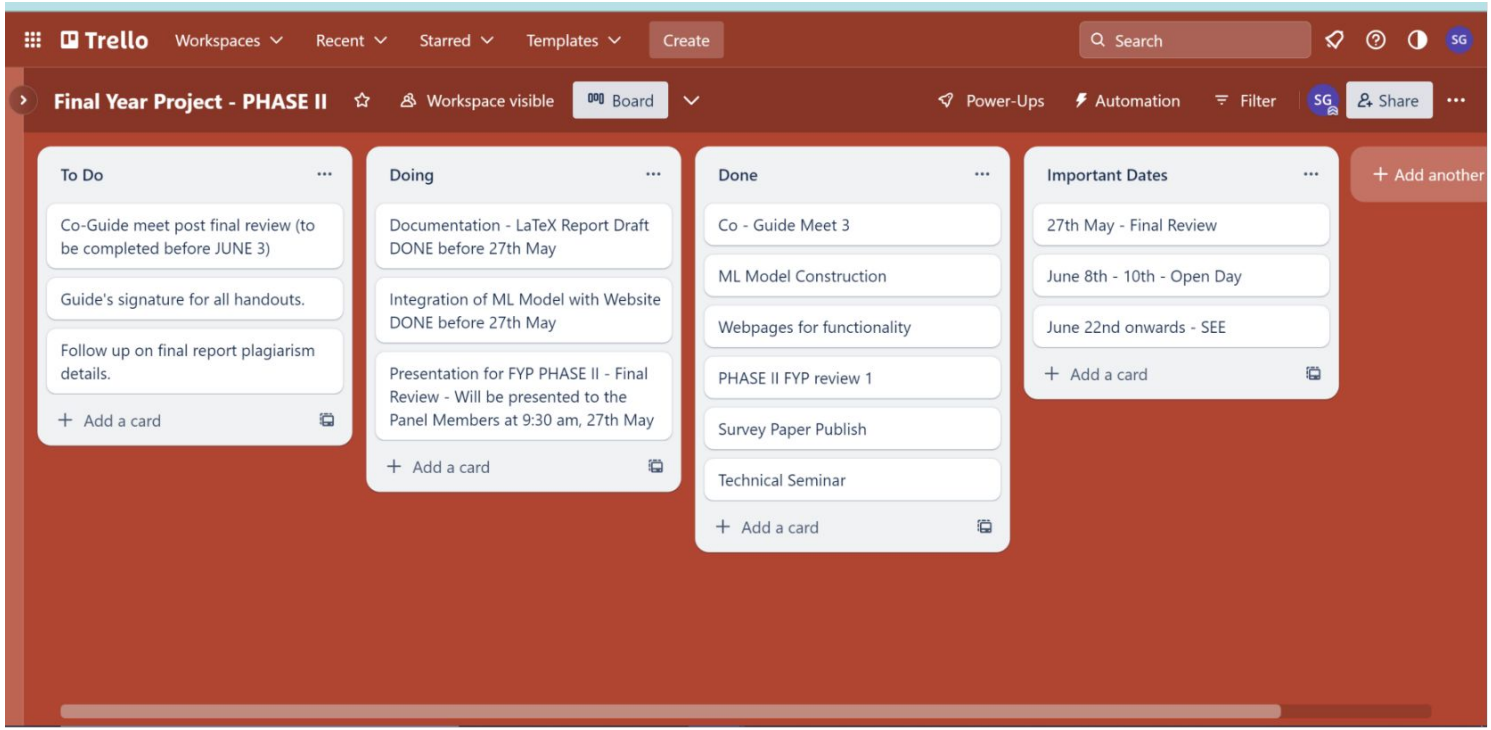


Dayananda Sagar Institutions

Department of Computer Science Engineering, DSCE

Atlassian Trello, a powerful project management tool, seamlessly integrates into our ML project, enhancing collaboration and organization. With Trello, we can create boards to map out the various stages of our ML workflow, from data preprocessing to model training and evaluation. Each task and subtask can be represented as cards, allowing team members to assign and track responsibilities, set deadlines, and monitor progress in real-time. The ability to attach files, share comments, and tag team members ensures seamless communication and collaboration throughout the project. Trello's intuitive interface and customizable features enable us to tailor the boards to our ML project's specific needs, providing clarity, structure, and transparency to the entire team. By leveraging Trello's robust project management capabilities, we streamline our ML project's execution, foster efficient teamwork, and ensure that every step is well-organized and executed with precision.

USAGE OF PM TOOL - ATLASSIAN TRELLO



Project Design:

Data Collection and Preprocessing:

We will leverage the extensive dataset provided by Crunchbase, extracting relevant startup information, including the Category List, Total Funding raised, country of operation, number of funding rounds, and funding dates.
The data will undergo thorough preprocessing, including cleaning, handling missing values, and encoding categorical variables, ensuring the data is in a suitable format for ML model training.

Feature Engineering:

In addition to the provided input parameters, we may explore additional feature engineering techniques to enhance the prediction model's accuracy. This could involve deriving new features based on the existing data or incorporating external data sources that may provide valuable insights into startup success factors.

ML Model Selection and Training:

We will experiment with various ML techniques, including but not limited to Random Forest, Logistic Regression, Gradient Boosting, and Support Vector Machines, to determine the most effective approach for our startup success prediction model.
We have identified the Random Forest ensemble technique as the optimal choice based on rigorous evaluation and comparison due to its ability to handle complex relationships, handle high-dimensional data, and provide robust predictions.

Model Training and Evaluation:

The selected ML model, Random Forest, will be trained on a carefully curated dataset derived from the Crunchbase data, with a focus on the input parameters mentioned earlier.
We will employ appropriate validation techniques such as cross-validation and performance metrics like accuracy, precision, recall, and F1 score to evaluate the model's performance and fine-tune its hyperparameters for optimal results.

Pickling and Model Integration:

Once the Random Forest model is trained and evaluated, we will pickle the model to store it in a serialized format, ensuring its seamless integration with the website and enabling fast loading for real-time predictions.
The pickled model will serve as the core component of our ML integration, allowing users to input startup details via the website's user interface.

Website Integration and API Calls:

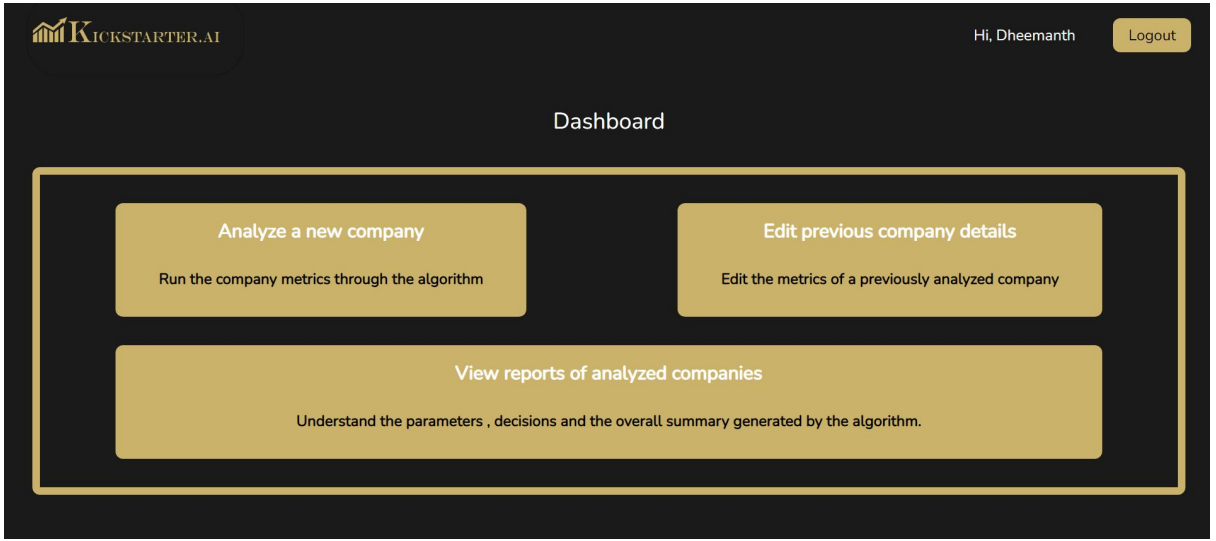
We will develop a user-friendly website that interacts with the ML model through API calls, providing an intuitive interface for users to input startup details and retrieve predictions.
The website will handle the data input validation, ensuring that all required parameters are provided and in the correct format before making the API call to load the pickled ML model.

Prediction Generation and Result Display:

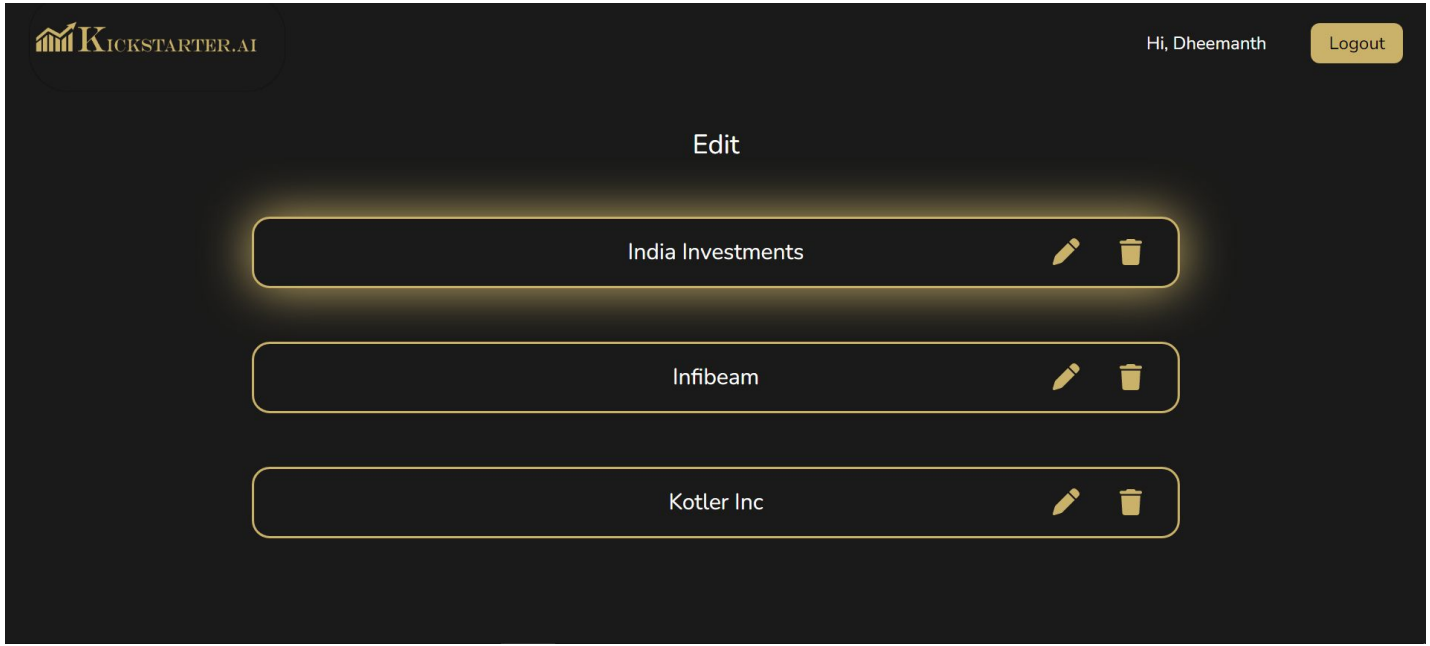
Upon receiving user input, the website will pass the relevant parameters to the pickled ML model via API calls. The ML model will process the input using techniques such as vectorization and serialization, generating a prediction regarding the startup's potential for success, including the likelihood of IPO transition or acquisition by a larger company.
The prediction results will be displayed on the website, providing users with valuable insights into the future prospects of their startups.

UI SCREENSHOTS

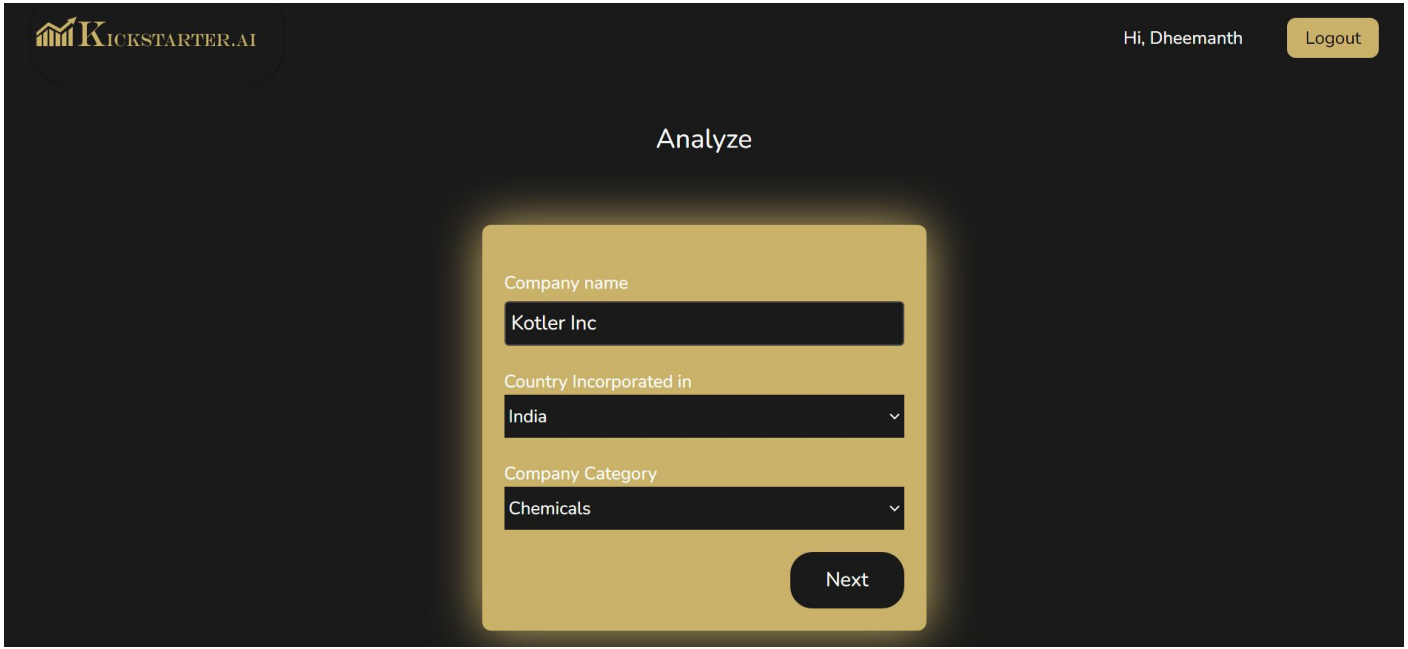
Dashboard:



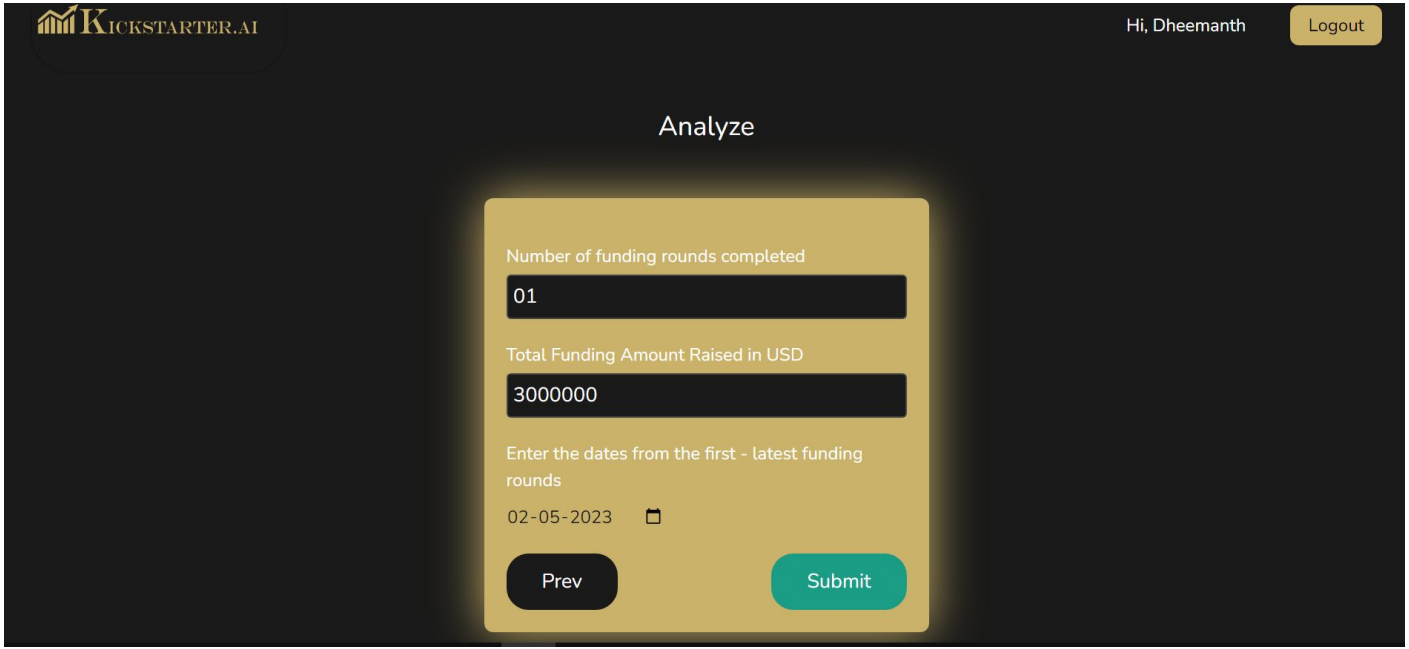
Edit Response:



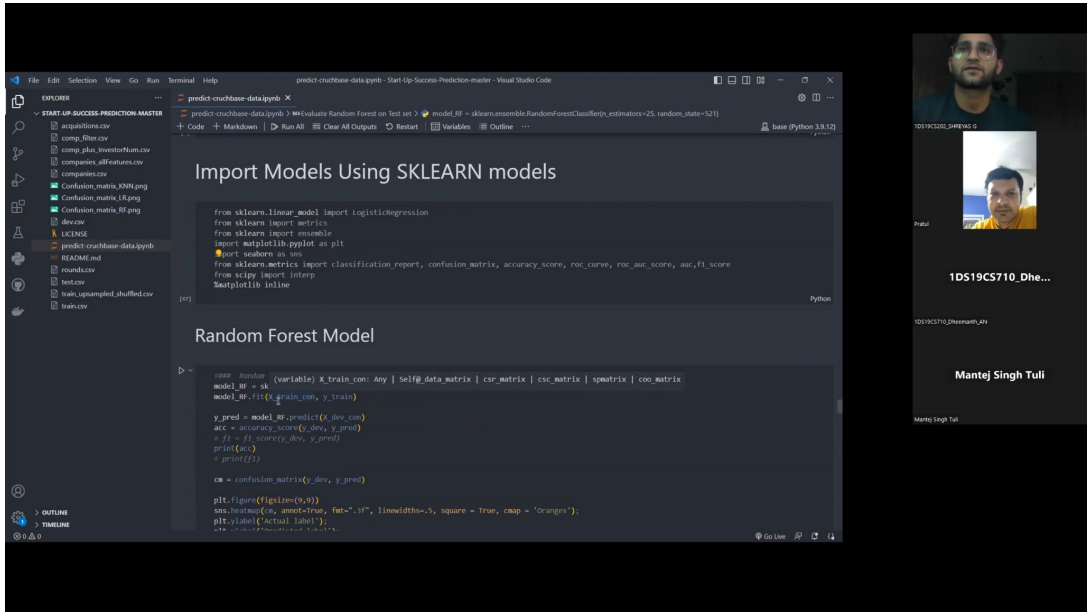
Analyze pt.1


KICKSTARTER.AI
Hi, Dheemanth
Logout
Analyze
Company name
Kotler Inc
Country Incorporated in
India
Company Category
Chemicals
Next

Analyze pt.2


KICKSTARTER.AI
Hi, Dheemanth
Logout
Analyze
Number of funding rounds completed
01
Total Funding Amount Raised in USD
3000000
Enter the dates from the first - latest funding rounds
02-05-2023
Prev
Submit

Co-Guide Meeting Outcomes:

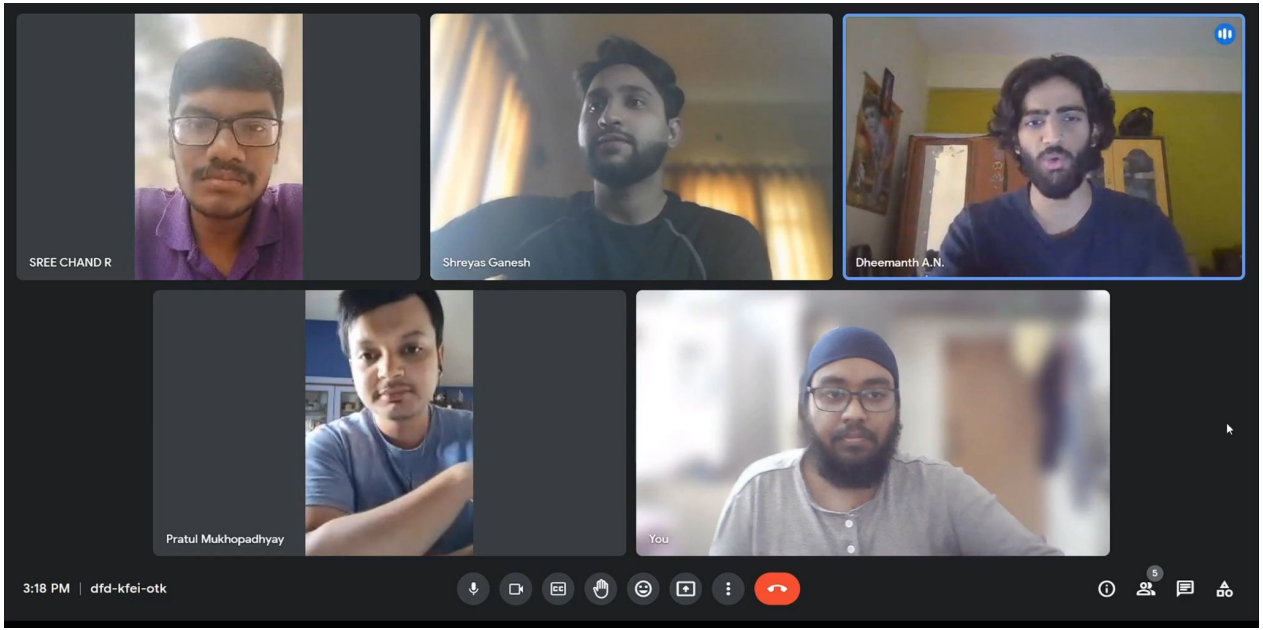


During our meeting with our co-guide, Pratul Mukhopadyay, held on April 11th, 2023, several significant suggestions were provided to enhance the project's quality and impact. One prominent recommendation was to publish a survey paper in a high-impact-factor journal. Following this guidance, we successfully identified and published our survey paper in the prestigious International Journal of Scientific Research in Engineering and Management (IJSREM).

Furthermore, Pratul emphasized the criticality of meticulously selecting input parameters with utmost certainty and conducting data preprocessing accordingly. Considering his expert advice, we carefully evaluated various factors and determined Category List, Total Funding raised by the organization, country of operation, number of funding rounds, and All funding dates as our precise input parameters. This strategic decision ensures that our model receives the most relevant and essential information for accurate predictions.

Moreover, during our meeting, Pratul emphasized the importance of providing a detailed report that not only delivers the prediction but also offers insightful explanations regarding the underlying reasons for each prediction. While we haven't completed this task yet, we have taken his suggestion into serious consideration and it will be an integral part of our future work. We are fully committed to providing comprehensive reports that delve into the various factors, patterns, and relationships uncovered by our model during the prediction process. By incorporating these explanations, we aim to make our predictions transparent and explainable, enabling stakeholders to gain valuable insights into the decision-making process.

Overall, the co-guide meeting with Pratul Mukhopadyay was an invaluable milestone in our project, where we received expert guidance on publication strategies, input parameter selection, and the significance of providing detailed reports with insightful explanations. Incorporating these technical recommendations ensures the robustness, credibility, and impact of our project.



CONCLUSION

In conclusion, our project proposes a comprehensive solution for predicting startup success by harnessing Crunchbase data and employing advanced ML techniques. By developing a user-friendly website integrated with our pickled ML model, we empower entrepreneurs and investors to make informed decisions based on data-driven predictions, ultimately increasing the chances of startup success and facilitating growth within the entrepreneurial ecosystem.