

Advanced Management Programme in Business Analytics Programme

Capstone Project Charter

The Capstone Project Charter is a formal agreement between the project sponsor and the student Team working on the Capstone project which clearly outlines the nature, extent, scope, approach and deliverables of the project.

The team is expected to deliver on all the commitments made in the project charter while the client is urged to actively participate in the creation of the project charter which then helps ensure that all stakeholders have consistent expectations from the project and agree on the approach and deliverables for the project.

Once the project charter is finalized, no deviations can be undertaken by any side without consultation and approval of AMPBA program office.

Each team will prepare a project charter as per the template given below.

Please finalize this document in consultation with the sponsor and it is mandatory to obtain their sign-off on this project charter by the deadline mentioned below:

Deadline for Project Charter – 14th Jan 2022

It is mandatory for the team to interact with both the sponsor at least once, to discuss the charter, before you finalize it.

It is mandatory to fill in every section of this form.

Please keep the project charter approach and deliverables consistent with the project description provided by the sponsor and his requirements for the project.

If the scope, deliverable, or schedule requirements of the project change significantly, a 'Change Request' should be submitted by the sponsor to the AMPBA program office. This request must be approved by us, in consultation with the faculty, to protect teams from scope creep during the project.

PROJECT DESCRIPTION

Name of the Client	Lotus Dew Wealth and Investment Advisors Pvt Ltd
Project Title	Stock IPO Scoring using NLP and AI
Project Team	<ol style="list-style-type: none">1. Chandra Shekhar Bhakat2. John Daniel3. Gauri Mayekar4. Shankarshan Mukherjee
Client Mentor	Mr. Abhishek Banerjee (CEO Lotus Dew)
Faculty Advisor	Dr. Vasudeva Varma (Faculty at ISB)
Project Team contact Details	Chandra Shekhar Bhakat – 9030164082 chandrashekhar_bhakat_ampba2021s@isb.edu John Daniel – 9811451592 john_daniel_ampba2021s@isb.edu Gauri Mayekar – 9833930997 gauriyogeish_mayekar_ampba2021s@isb.edu Shankarshan Mukherjee – 9700012321 shankarshan_mukherjee_ampba2021s@isb.edu
Project Mentor Contact Details	Mobile: +91 77310 82221 Official Email: abhishek.banerjee@lotusdew.in
Faculty Advisor Contact Details	Prof. Vasudeva Varma Email: vv@iiit.ac.in
Project Description (as per PD form)	<p>IPO prospectus has 300-400 pages of information. For an average retail investor, its very difficult to follow this information and compare against peers. The project aims to make this systematic by extracting the required information from IPO prospectus. In the first phase of the project, you will do review of academic literature to define the aspects that define post IPO performance and signify a good quality issue. In the second phase you will write code to automatically process IPO prospectus to extract these parameters and annotate the PDF by highlighting the area that was used to extract the information for tracking.</p> <p>Overall the aim of this project is to setup basic NLP, Entity Recognition and parsing technology and define analytics around peer group analysis including foreign peers. For this project, we will specifically focus on relatively unknown but upcoming SME IPO segment. The key difference is that these are fully underwritten</p>

	<p>instead of main board listing which is best effort raise by the investment banker.</p> <p>Factors that we would look for are promoters, related party disclosures, litigations, industry dynamics, peers, past performance by its bankers etc.</p>
--	---

PROJECT PLANNING

Business Issue:

- Business Context

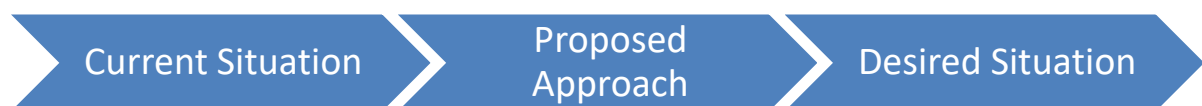
It is difficult for an investor to read through a 300+ page IPO document to understand about the company and do a peer comparison. This results in investors' not having the adequate know how of the company's performance which in turn might result in investors staying away from IPOs or making uninformed bets.

Further, today a lot of financial details are churned to assess the performance of a company, but the non-financial data is not taken much into account. There could be some critical insights in the non-financial data which could enable the investor make an informed decision.

- Area of impact

The investors will be guided towards better / more informed and unbiased decision making.

Need Gap Situation :



Current Situation:

- The investor either relies on the insights / forecasts of various securities agencies to decide on the investment Or analyses the documents him/herself and does a peer comparison.
- There is no module / application with the sponsor which can automatically parse an IPO document and extract critical information. Also the metrics for scoring an IPO document on the non financial parameters are not available.

Proposed Approach:

- Do an extensive secondary research to understand the critical non financial parameters that defines success of an IPO. Identify research papers and other documents to develop this understanding and come up with metrics to score a document
- Prepare train and test data set.

- Extract relevant data and perform EDA
- Prepare a model to score the document basis the details extracted. Test on the test data and make it better

Desired Situation

- An IPO document fed by the user gets auto scanned and gives a score to the IPO

Proposed Project Work Flow:

Define metrics to be measured / monitored

- Access sources such as Google scholar, Ebsco (LMS resource) and internet in general to identify research work done in this domain.
- Interact with ISB faculties to get their insights on the topic as well
- Define a criteria in discussion with the sponsor and the guide for choosing any metric for assessment
- Finalize the metrics iteratively

Define Success / Failure criteria of an IPO

- Secondary research for defining successful IPOs
- Study the performance of IPOs in the last 5 years
- Analyze data to identify patterns / groups
- Finalize the criteria

Code to extract data from documents

- Parse the file and read relevant data
- Use existing libraries and customized codes wherever needed
- Format and clean the data extracted for further use
- Feature engineering

Build and test model

Evaluate the model and deploy

Data Plan and Software requirement for each deliverable:

Data requirements :

1. Access to Refinitiv database of more than 20000 companies. This data includes financial results, price data and also peer data
2. IPO prospectus from the financial sites. (200 documents will be provided by the sponsor)

Data collection:

Some of prospectus documents would be fetched by ourselves and some would be provided by the sponsor.

Data security:

We will ensure that no part of the code will be shared publicly as well as try to use tools like Google Colab so that the code is available in one place. While the hosting can be anywhere we will try to store the trained model in a pickle file or an API to reuse it.

Software used:

We will be using Python to parse the PDF files, extract named entities and create the model.

Deliverable Project Management:

Deliverable	Start Date	Deadline	Key Analysis Parameters	Data Sources	Outcome expected	Team Member Responsible
Literature review of factors linked to IPO performance - Identify research paper with high number of citations - Check for validity of the findings of research paper - Create a repository of research data with necessary comments - In person discussions with ISB faculties to gather their insights	<i>20th Jan</i>	<i>30th Jan</i>	<i>Adequate references to support the metrics identified</i>	<i>Google scholar, Ebsco</i>	<i>Metrics identified and finalized with Guide and Sponsor</i>	<i>Chandra Shekhar</i>
Define Success / Failure criteria of an IPO - Collect IPO data for the last 5-7 years - Analyse the performance of data from the date of listing to different time horizons - Categorize the performance and define the classes	<i>20th Jan</i>	<i>5th Feb</i>	<i>Measurable and objective criteria</i>	<i>Google scholar, In person references</i>	<i>Success criteria on an IPO is defined</i>	<i>Shankarshan</i>
Code to extract data from documents - Parse pdf to extract data - tables, text etc - Use NLP techniques to scan document from the table of contents - Create a structure data in terms of flat file - Use NLP techniques to score the qualitative data - Perform feature engineering	<i>20th Jan</i>	<i>20th Feb</i>	<i>Efficient with scope for easy scale up</i>	<i>IPO documents</i>	<i>Executable code to parse the files</i>	<i>John Daniel</i>
Build and test Module - Define a scoring method - Score the data basis the scoring method - Test against the labelled data and refine iteratively	<i>20th Feb</i>	<i>10th Mar</i>	<i>Accuracy</i>	<i>IPO documents</i>	<i>Model with expected accuracy</i>	<i>Gauri</i>

Deploy - Test the module on real time data - Documentation	<i>10th Mar</i>	<i>20th Mar</i>	<i>Efficiency</i>	-		<i>Gauri</i>
---	-----------------	-----------------	-------------------	---	--	--------------

Progress Updates

The team will send update emails every week to the sponsor, faculty mentor and AMPBA program office. Each update will contain a brief summary of progress during the time and key insights that have been revealed. Other key milestones include:

Project Charter Finalization	19 th Jan 2022
Weekly Reports	Every Sunday
Mid Review	19 th & 20 th Feb 2022
Final Report Submission	3 rd April 2022

Agreement

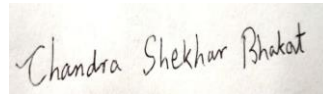
This charter and accompanying deliverables have been mutually discussed and agreed by:

Abhishek Banerjee
Project Mentor Name (Sponsor)

DS LOTUSDEW WEALTH AND INVESTMENT ADVISORS PRIVATE LIMITED 1
Digitally signed by DS LOTUSDEW WEALTH AND INVESTMENT ADVISORS PRIVATE LIMITED 1
Date: 2022.01.19 21:30:22 +05'30'

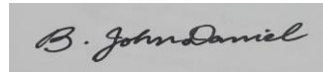
Project Mentor Signature (Sponsor)

Chandra Shekhar Bhakat
Team Member #1 Name



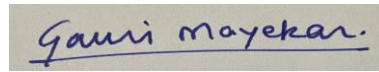
Team Member #1 Signature

John Daniel
Team Member #2 Name



Team Member #2 Signature

Gauri Mayekar
Team Member #3 Name



Team Member #3 Signature

Shankarshan Mukherjee
Team Member #4 Name



Team Member #4 Signature

Capstone Scope Change Request

Name	Email Address
Company	Telephone Number(s)

Please describe the nature of the change requested:

Approximately how many additional hours of effort will this change require?

What other delays, costs, or risks would this change introduce?

What part of remaining project work can be forgone to accommodate this change?

Please email this form to hemanth_kumar@isb.edu

Decision will be conveyed within one week.

Student Team Recommendation:

Faculty Recommendations:

Other Considerations:

☐ Accepted

☐ Rejected