



IIT KHARAGPUR

INTER IIT TECH MEET 10.0

25-27TH MARCH 2022

DIGITAL ALPHA'S SEC FILING ANALYZER FOR SAAS COMPANIES

The SEC's EDGAR database contains terabytes of documents and data, including press releases, annual corporate filings, executive employment agreements, and investment company holdings. While EDGAR has existed for over twenty years, scholars have had difficulty conducting or reproducing research based on EDGAR data. Researchers often spend a lot of time and money developing and redeveloping code to retrieve and parse EDGAR data with no common bottom-up framework.

PROBLEM STATEMENT

CONTEXT

SaaS companies are customer-driven and are heavily dependent on their customer base. There are a set of metrics that can showcase the health of the SaaS companies and their aspects. These metrics and numbers are not readily available on publicly reported SEC Filings and need to be chalked out from the available forms (10-K, 10-Q, 8-K, etc.). Accessing these metrics can be of tremendous value to the right users.

Forms that are being filed are very detailed and take time to analyze. For analyzing a form, the primary task is chalking out the key SaaS numbers. There is a need to have simplified dashboards to have a quick overview of Key SaaS Goals.

FORMS FOR SAAS COMPANIES

10K: It is an annual report of the company. Footnotes in 10K forms can help to flag sketchy accounting practices.

10Q: It is the company's quarterly report that reflects changes in inventory and working capital. It can be used to draw comparisons between companies.

8K: Used for important announcements, like executive appointments, etc., between filings of a majority of forms - investors will get non-financial inputs, for example, when TATA sons had a management change or similar situations.

PROBLEM STATEMENT

- Create an interactive dashboard capable of analyzing SEC Filings of all US Listed SaaS companies for all the available years of data.
- The dashboard should be a Key SaaS goal analyzer for the user/client.
- With this dashboard, one should be able to pull up information and related statistics about any SaaS company they are interested in and extract useful information with ease, rather than reading voluminous filings that are often 50–100 thousand words in length.

SKILLS REQUIRED

Web Development & Deployment, Web Scraping, Machine Learning, UI/UX

Note that you are responsible for complying with EDGAR's access terms and conditions.

GUIDELINES

- The dashboard should have a search bar to enter the required query by the user - company name, date range, metrics, type, etc.
- Key SaaS goal analyzer can be based on growth, profitability, potential, etc.
- The dashboard should showcase the drill-down map of the Key SaaS goals of the searched company for the specified date range.
- For unavailable data, the participating team should choose appropriate alternative resources.
- Table of Content: Break down the lengthy SEC filings into separate sections, each of which deals with different aspects of a company's reporting, making it easy to navigate.
- ML model to be trained and implemented in the web tool for any chosen usage - Example: Analyze the Annual report for the given time frame to get the sentiment from the SEC filings to assess the company's situation.
- Come up with a single call API that creates a dataset in a few lines of code for any duration of time and numerous tickers. The extracted data frame must be written as a CSV file and downloaded.
- Graphs and charts to showcase the trends and changes in metrics for a given date range.
- Metrics classification for SaaS companies- Good, bad, or neutral for each drill-down tab.
- Competitor comparison page - Dashboard with a feature to compare two or more companies for the same date range with trend graph overlay and key SaaS goal drill-down overlay.
- The web tool must be hosted using standard hosting solutions and databases only.

EVALUATION

- Identification of Problem Statement and Solution architecture
- Dashboard features - UI/UX, Ease of use, speed of execution of the drill-down overlay
- Machine Learning model efficiency and analysis
- Random test case passing (6 test cases) - Proper analysis of SaaS metrics
- Correct analysis of the key SaaS metrics and metrics classification
- Accuracy of metrics classification for SaaS companies and accuracy of generated sentiments; compared with actual data in forms and filings
- Scalability and compatibility of web tool
- Innovation in dashboard features
- Presentation skills

Brownie points

- Extra points will be rewarded for including additional features and seeing the overall inclusivity of the web tool.
- Show pros and cons of a company on the dashboard (somewhat like Ticker tape) - Smart Narrative Note that you are responsible for complying with EDGAR's access terms and conditions.

SUBMISSIONS

The evaluation is to be conducted in 2 steps: (400 points)

- Report and Code Submission - 200 Points
- Final Presentation - 200 Points

REPORT AND CODE SUBMISSION - 22 MARCH 2022

- Github repository with well-defined documentation files
- Report: Detailed web tool architecture, problem approach, key insights
- Machine Learning Model Performance - Precision, Recall, and F1 Score

An overall score would be computed which would be a simple average of the above three metrics.

The report needs to be submitted in PDF format. It should not exceed 15 pages (excluding the Cover Page).

FINAL PRESENTATION - 26 MARCH 2022

- All participating teams would be required to give presentations of their proposed solution on 26th March.

Please note that the Final presentation should wrap up within 15 minutes which will be followed by Q&A (5 min).

Send your submissions at this email: submissions@interiit-tech.org

List of target SaaS companies: <https://bit.ly/3ByDLc3>

REFERENCES

- <https://www.entrepreneurs.com/saas-metrics/>
- https://edgar-online.com/data/EDGARPro_Brochure.pdf
- <https://www.investopedia.com/articles/fundamental-analysis/08/sec-forms.asp>

Team size for this event is maximum 10 participants.
Participation awards shall be awarded to all participants.