TresVista
Catalyzing Your Impact

Get a Competitive Edge Using the Power of Automation and AI

Automation and Data Analytics in PE and VCs

After years of investing in AI start-ups, Venture Capitalists are now utilizing the power of AI and data themselves to gain a competitive edge. According to Gartner, 75% of VCs will use machine learning and artificial intelligence in their decision-making process by 2025. However, the adoption has been far slower than expected. The majority of VCs continue to operate their workflow in the traditional way while betting on disruptions in other industries. A few VCs have taken the lead here and paved the path for others, but the use of machine learning and artificial intelligence is still in the early stages. There is also a big difference in the approach followed as some firms have chosen to be public about their data-driven approach (EQT Ventures, Hone Capital), while others have been reserved in their communication with the masses.

Below, I discuss some ways that VC and PE firms are utilizing machine learning and AI to automate processes.

1. Sourcing opportunities

For sourcing new opportunities, VCs and PE firms can implement automation to make the process more efficient, increase the number of opportunities coming to their funnel, and have more information about the market. Early-stage VCs can scrape and track the websites of incubators to track the latter's portfolios. Late-stage or growth VCs can automate tracking the portfolio of early-stage VCs as well as their competitors. Similarly, PE funds can automate pulling information from exhibitions and conferences where many hundreds of potential target companies participate, significantly reducing the time to collate such data and allowing more time for analysts to spend on more value-added activities. Using automation also reduces the chances of mistakes and data duplication.

Early-stage investors would like to track multiple sources of information to ensure they are aware of an opportunity before competitors. They can use the power of automation to track specific angel investors, key personnel movement, social networks like LinkedIn, public registers, databases like Crunchbase, specialized media like TechCrunch, patent registration, or any other industry-specific sources at a large scale. Similarly, VCs or PE firms can track website traffic, app downloads, social media noise, influencers, social media sentiment, etc. to keep abreast with up-and-coming companies or brands.

2. Screening

VCs are also using data analytics on screening companies to support the processing of the increased inflow of deals. By using machine learning on large databases of historical information to score companies based on multiple data points, they are able to shortlist companies. This has allowed for quicker review processes, especially for companies that will not make the cut. This also allows the investment teams to complete assessments faster. Certain parameters that can help predict the probability of success and fit for asset managers are the background of the team, the team mix, total money raised, syndicate lead's area of expertise, market size, product segment, competitive landscape, business model, etc. Using these, a scorecard can be built to evaluate companies, indicating their chance of success. This allows selection to be automated to an extent, based on a set of parameters.

EQT Ventures, Correlation Ventures, and Hone Capital are some of the VCs implementing a data-driven approach to screening opportunities. These firms have built and trained machine

learning algorithms using historical data pulled from various databases and using them to reduce selection timelines.

3. Benchmarking and tracking

PE and VC firms can automate benchmarking and tracking of their competitor firms. Opus Capital, an early-stage VC firm, is using a simple data model to keep track of the ecosystem by shadowing 25-30 peer firms. Tracking data points like new investments, follow-on investments, sector niche, geography, business models, etc. on a quarterly basis, the firm keeps a tab on the VC ecosystem. This allows the firm to track the ecosystem closely, catch any changes in trends in terms of sectors, business models, and geography relatively early, and better understand co-investment scenarios. Data can be collated from multiple sources like Pitchbook, Crunchbase, VC company websites, VC/PE-specific news websites, etc., for the chosen list of firms.

4. Investor reporting and portal

Investor reporting is one area where VCs and PE firms can work to improve reporting and investor experience. Digitalizing and allowing the investor to view the reports on demand and from anywhere significantly improves customer communication and reduces manual effort. The online investor portal allows for stringent access control based on login profile. Interactive visualization also allows the investor to analyze the data provided.

5. Automated portfolio monitoring and visualization

One area that takes a lot of manual effort for VCs and PE firms is portfolio monitoring. Portfolio companies share data via email on a monthly, quarterly, and annual basis, which is manually taken and input in Excel or presentation format for everyone to consume; from analysts to senior management. Then there might be instances where the management asks for additional data, which would again require more effort and time to turn around. This process can be automated, where all information can be picked up automatically on a periodic basis and showcased in an interactive dashboard, which can be accessed from anywhere. It also allows for customization like moving between GAAP and management reported numbers, looking into nuances of certain KPIs, etc. This improves the management's consumption of the information, allows the user to dig deeper into the data as required, provides access from anywhere, and reduces manual effort.

All in all, bringing automation into the equation can help reduce the burdens of asset management to a great extent. TresVista's Data Intelligence Group (DIG) has worked with several private and public asset managers to help automate various processes. Our full-stack data team with data scientists, data engineers, and data analysts helps clients embrace a data-driven approach to decision-making and automate processes to drive efficiency.

Case Studies

Portfolio Reporting Architecture Plan

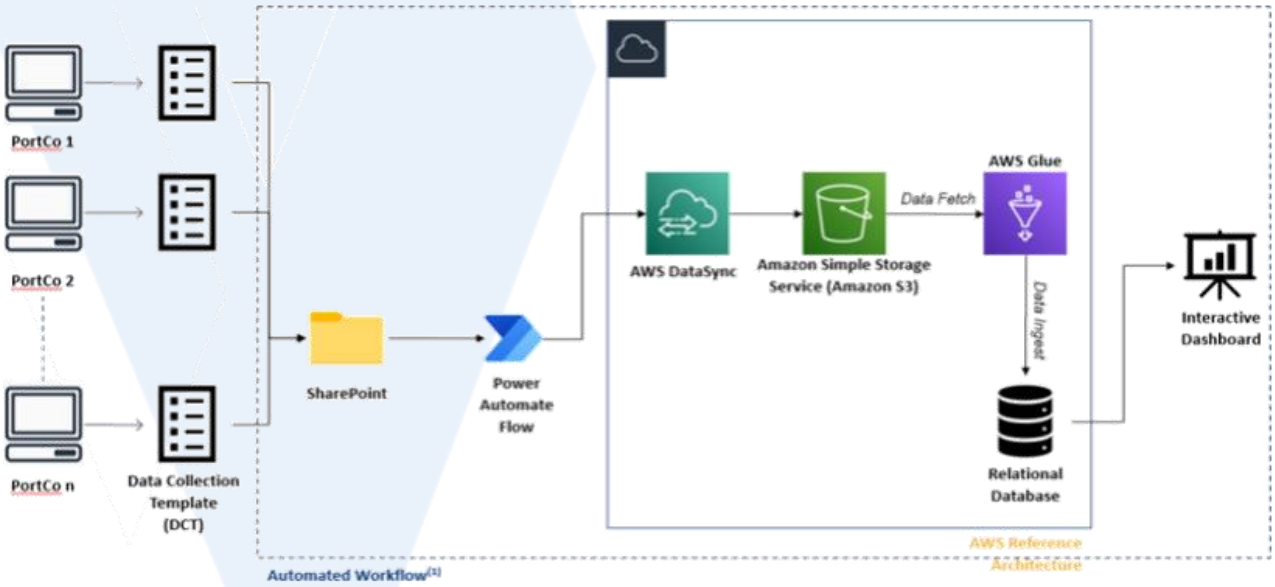
Data Sourcing, Integration, and Manipulation

Request and Guidelines Provided

- To collate the financial and operational KPIs for Portfolio Companies to design a standardized report
- To automate the report generation process with limited manual intervention
- To design end-to-end solution from data sourcing, integration, manipulation to report generation
  - No secondary off-the-shelf Portfolio Reporting software is used

Methodology and Final Deliverable

- The team designed an Excel based Data Collection Template (DCT) for the stakeholders to add the financial and operational metrics for each of the portfolio companies
- The stakeholders must populate the DCT based on the existing models that are received from each of the portfolio companies
- Once filled, the user can upload the template to the SharePoint location with a click of a button, from where the data will be automatically pushed to the database
- In case there are any issues with the uploaded model, the stakeholder would receive an auto generated detailed email with a list of the issues
- The Power BI/Tableau dashboard would leverage the database to display a standard report



Tools/Technology used: VBA, Microsoft Office 365, Power BI / Tableau.

Data Engineering

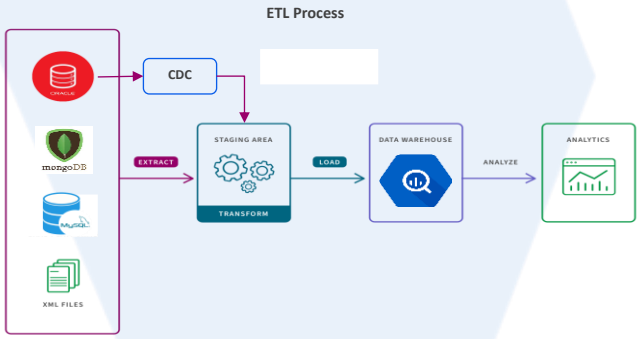
Data Sourcing, Screening, Loading, and Visualization

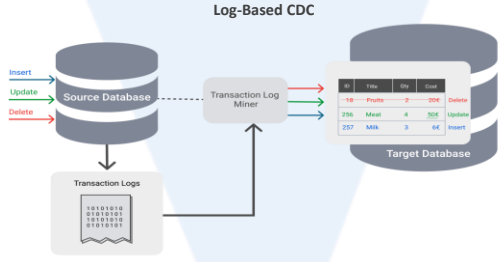
Request and Guidelines Provided

- To design a Data Warehouse to capture data across multiple sources like MySQL, Oracle, and Mongo DB into BigQuery
  - Oracle being the main database which is used by the entire firm, the ETL process should not impact the database performance
- The data from BigQuery would be further used by the reporting and analytics team to generate insights and share them with the stakeholders

Methodology and Final Deliverable

- Data models across the databases were studied (MySQL had company/employee data, Oracle had tick-level accounting data, and MongoDB had ticker-specific reviews & newsletters), and BigQuery schema was finalized
- PySpark-based ETL scripts were built for MySQL and MongoDB, and log-based CDC scripts were built and scheduled for periodic updates
  - Log-based CDC was a highly efficient approach that limited the impact on Oracle DB with minimal/zero-downtime
- Multiple data and business-specific controls were added in the ETL scripts to maintain the quality and integrity of the data moving into BigQuery





Tools/Technology used: SQL, PySpark, BigQuery

Data Engineering

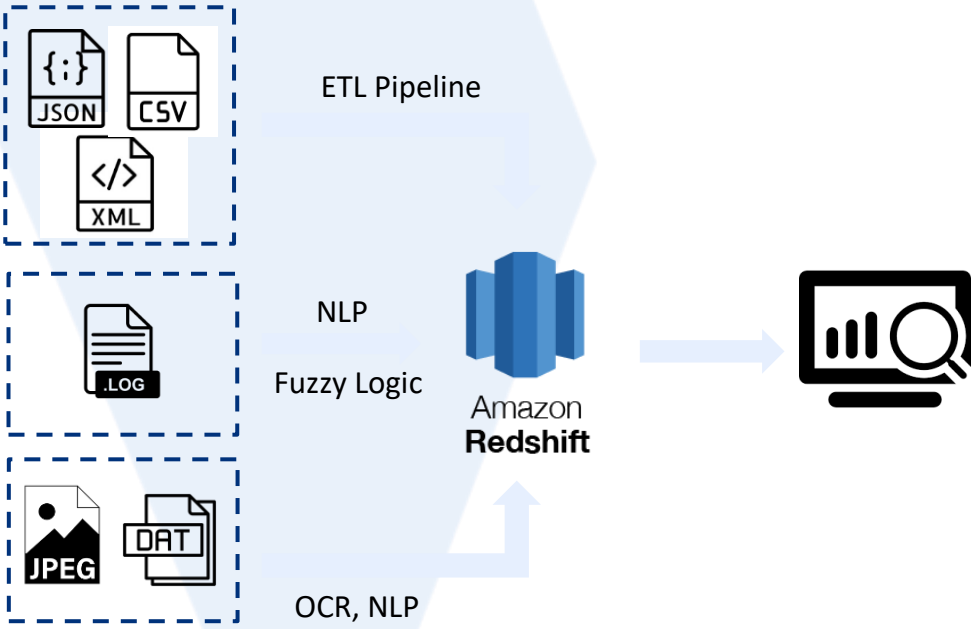
Data Collation and Standardization Across Multiple Sources

Request and Guidelines Provided

- To create a centralized data warehouse for a private equity client by,
  - Extracting data from different sources like flat-files, logs, images, etc.
  - Transforming the data into a standardized format based on business logics
  - Storing it for in-depth analysis

Methodology and Final Deliverable

- Extracted data of different types, such as CSV, .log files, PDFs etc., from multiple data sources like emails, deal documents, meeting logs etc.
- Performed data transformation process using a combination of business logic, NLP, Fuzzy Logic, and OCR to extract specific deal details, sentiment of the meetings, text from images etc.
- Implemented a set of rule-based algorithms to structure the data based on specific patterns and created automated pipelines that integrates with the database management system to load the transformed data into Redshift
- Performed data analysis to gather insights to make data-driven decisions using visualization tools



Tools/Technology used: SQL, PySpark, Redshift

To see a sample interactive dashboard please click the below link:

https://app.powerbi.com/view?r=eyJrIjoiNDdmMGIxZTAtZjg0Yi00NjVmLTgwYTMtNmZiYWFmZTA3YzE4IiwidCI6IjQyNTgzMzJjLTRhN2YtNDdjYy04ODFhLWNkMGU0Y2VhN2Q5ZiJ9&embedImagePlaceholder=true

DISCLAIMER:

This document is provided for information purposes only. The information is believed to be reliable, but TresVista does not warrant its completeness or accuracy. It should not be used, relied upon, or treated as a substitute for any professional advice. Opinions, estimates, and assumptions constitute our judgment as of the date hereof and are subject to change without notice. TresVista may, at its absolute discretion and without any obligation to do so, update, amend or supplement this document. TresVista disclaims any and all liability arising from actions taken in response to this report. TresVista, its employees, and any persons associated with the preparation of this report are in no way responsible for any errors or omissions in the report resulting from any inaccuracy, misdescription or incompleteness of the content provided. This material is not intended as an offer or solicitation for the purchase or sale of any financial instrument.

salessupport@tresvista | www.tresvista.com