Refract: Empowering Enterprise AI/ML Adoption

# Introduction

Refract is a self-served, unified data analytics platform that abstracts the operationalization of the AI lifecycle, behind the scene. It paves a faster and smoother way to build, train, deploy, search, manage and monitor the enterprise AI/ML model portfolio at scale. It is aimed at enabling the enterprises to shift from small scale experiments in silos to collaborative enterprise scale production. The platform facilitates the AI everywhere paradigm for the enterprises and allows for massive adoption of AI across the enterprise. Refract can be used by Data Scientists and data driven organisations to democratize data science in an enterprise. The solution helps in taking care of deployment, scaling and versioning of AI models, so that data scientists can focus on creating AI models without worrying about the deployments and so on. Fosfor by LTI is committed to client centricity. With Fosfor, you can accelerate your data-to-decisions journey, adopt to cutting-edge data technology, and achieve your strategic goals seamlessly. The Fosfor Product Suite provides business users with intuitive access to data lifecycle tools without requiring IT intervention. Moreover, it provides IT and data teams with tools to optimize every operational aspect of conducting advanced analytics across all relevant data sources. Refract can help the bank by providing a platform that enables data scientists to configure custom environments for themselves by adding the required support libraries, that are preconfigured in notebook image when it is spawned. It enables data scientists to create isolated development environment in Kubernetes cluster. It enables data scientists to schedule notebooks which help them to get results on timely basis. It enables sharing of projects allowing access to notebooks created and shared by you. It enables you to install required libraries from a drop down menu. It enables you to launch different notebook environments on the click of a mouse. It enables data scientists to track their work (using integration with Git).  
  
Question: What is the purpose of a project in Refract? How do you create a project?  
Helpful Answer: A project is the starting point to begin work in Refract. It is a collection of datasets, notebooks and models. It provides a single view to keep track of resources. You can create a project on your own and invite others to collaborate / review them. To create a project, you need to follow these steps:  
1. On logging in to the Refract platform, the Project page, listing the existing projects is displayed.  
2. On the top-left corner, click on Add New Project. The Add Project window is displayed.  
3. In the Project Name and Project Description fields, enter a name and description for the project. The project can be created without adding a repository. You can either skip the option of adding repository by selecting No radio button to avoid unnecessary creation of repositories to the project, add an existing repository by selecting Yes or select the default repository by clicking on Default radio button.  
4. Click the Create button to create new project.  
  
Question: What is the purpose of the Default Settings tab in the Project home page? What are the different configurations that can be set in the Default Settings tab?  
Helpful Answer: The Default Settings tab in the Project home page displays multiple default configurations like configured Cluster, Spark Resource, Snowflake, Big Query, Parameters and Data Prep Settings. Parameters are basically the environment variables whose scope is limited to that particular project. Data Prep Settings are used to configure global write settings and strategy of wrangled (output) dataset in Data Prep. You can add new parameters from here. These parameters can be used in global parameters and in any schedules. The Default Settings tab is used to set the default configurations for the project. These configurations can be used in the notebooks and models created in the project. The different configurations that can be set in the Default Settings tab are:  
1. Configured Cluster  
2. Spark Resource  
3. Snowflake  
4. Big Query  
5. Parameters  
6. Data Prep Settings  
  
Question: What is the purpose of the Share tab in the Project home page? How do you share a project with other users?  
Helpful Answer: The Share tab in the Project home page displays the list of users with whom the project is shared with. You can Revoke and Share the project from here too. The Share tab is used to share the project with other users. To share a project with other users, you need to follow these steps:  
1. Click on the Share tab in the Project home page.  
2. Click on the Share button.  
3. Enter the email address of the user with whom you want to share the project.  
4. Select the access type for the user. The access type can be either Contributor or Owner.  
5. Click on the Share button to share the project with the user. The user will receive an email with the link to the project. The user can click on the link to access the project. If the user does not have an account on Refract, he/she will be prompted to create an account. Once the user creates an account, he/she can access the project. If the user already has an account on Refract, he/she can access the project directly.  
  
Question: What are the different services offered by Refract? What are the different languages supported by Refract?  
Helpful Answer: Refract offers the following services:  
1. Register Model  
2. Python  
3. R  
4. Pyspark  
5. Add Artefacts  
6. Python  
7. R  
8. Pyspark  
9. Deploy Model  
10. Python  
11. R  
12. Pyspark  
13. Load Model  
14. Python  
15. R  
16. Pyspark  
Refract supports the following languages:  
1. Python  
2. R  
3. Pyspark  
  
Question: What is the purpose of Refract platform? What are the high level features of Refract?  
Helpful Answer: Refract aims to optimize the end-to-end life cycle of Machine learning applications by simplifying various tasks of MLOPS like accessing the right data, training model, publishing and deploying models, model monitoring and evolution in a guided form using its capabilities. It minimizes efforts of creating model by using its extensive Automated ML feature without having much understanding of Data. It eases the life of Data scientist by providing multiple metrices to understand the model outcome. It provides model explanation to make users model trustworthy and understandable for novice users. Refract is a self-served, unified data analytics platform that abstracts the operationalization of the AI lifecycle, behind the scene. It paves a faster and smoother way to build, train, deploy, search, manage and monitor the enterprise AI/ML model portfolio at scale. It is aimed at enabling the enterprises to shift from small scale experiments in silos to collaborative enterprise scale production. The platform facilitates the AI everywhere paradigm for the enterprises and allows for massive adoption of AI across the enterprise. Refract can be used by Data Scientists and data driven organisations to democratize data science in an enterprise. The solution helps in taking care of deployment, scaling and versioning of AI models, so that data scientists can focus on creating AI models without worrying about the deployments and so on. The high level features of Refract are:  
1. Enables data scientists to configure custom environments for themselves by adding the required support libraries, that are preconfigured in notebook image when it is spawned.  
2. Enables data scientists to create isolated development environment in Kubernetes cluster.  
3. Enables data scientists to schedule notebooks which help them to get results on timely basis.  
4. Enables sharing of projects allowing access to notebooks created and shared by you.  
5. Enables you to install required libraries from a drop down menu.  
6. Enables you to launch different notebook environments on the click of a mouse.  
7. Enables data scientists to track their work (using integration with Git).<|im\_end|>

# Understanding of the problem statement

The bank is facing several challenges in building and maintaining their analytics infrastructure. The current infrastructure is not scalable and is not able to handle the increasing volume of data. The bank is also facing challenges in integrating data from different sources and in ensuring data quality. The current infrastructure is also not able to support real-time analytics, which is becoming increasingly important in today's business environment. The limitations of the current infrastructure are affecting the bank's ability to make timely and informed decisions, which is impacting their business. The bank needs to invest in a modern analytics infrastructure that is scalable, flexible, and able to support real-time analytics. This will enable the bank to make better decisions and stay competitive in today's business environment.<|im\_end|>

# Proposed Solution

Refract can help the bank overcome the challenges they are facing by providing a self-served, unified data analytics platform that abstracts the operationalization of the AI lifecycle, behind the scene. It paves a faster and smoother way to build, train, deploy, search, manage and monitor the enterprise AI/ML model portfolio at scale. It is aimed at enabling the enterprises to shift from small scale experiments in silos to collaborative enterprise scale production. The platform facilitates the AI everywhere paradigm for the enterprises and allows for massive adoption of AI across the enterprise. Refract can be used by Data Scientists and data driven organisations to democratize data science in an enterprise. The solution helps in taking care of deployment, scaling and versioning of AI models, so that data scientists can focus on creating AI models without worrying about the deployments and so on. Refract offers a number of features that make it a good fit for the bank's requirements, including the ability to configure custom environments, create isolated development environments, schedule notebooks, share projects, install required libraries from a drop down menu, launch different notebook environments on the click of a mouse, track work using integration with Git, and more. Additionally, Refract offers automated machine learning, which makes building models easier by running various feature engineering and model building techniques on the raw data. Refract automates the most tedious tasks of machine learning by trying out possible pipelines to find the best fit pipeline for the input data. Refract also offers data drift analysis, which provides an understanding of model decay, better optimizes hyperparameter and monitors shift in data trends. Finally, Refract offers the ability to bring your own container, which enables a web-based IDE in the platform to be loaded with minimal code changes like SAS studio, VS Code, Py-Charm and so on. You only need to provide the following details to spin up the IDE: container registry link, command to run container, port on which container runs, arguments required for container, and context path, if any. All of these features make Refract a good fit for the bank's requirements.<|im\_end|>

# Technical Approach

The text does not provide enough information to answer the question.<|im\_end|>

# Project Plan

Project Plan¶  
  
Project Name: Refract Implementation  
  
Project Manager: John Doe  
  
Project Start Date: 01/01/2022  
  
Project End Date: 01/01/2023  
  
Project Description: The project aims to implement Refract in the organization to streamline the data science process and improve the efficiency of the data science team.  
  
Key Milestones¶  
  
Milestone 1: Project Initiation  
  
- Project Charter  
- Project Plan  
- Project Kickoff Meeting  
  
Milestone 2: Refract Setup  
  
- Refract Installation  
- Refract Configuration  
- Refract Integration with GitLab  
  
Milestone 3: Data Ingestion  
  
- Data Ingestion from various sources  
- Data Cleaning and Preprocessing  
- Data Transformation  
  
Milestone 4: Model Development  
  
- Model Development using Jupyter Notebooks  
- Model Training and Evaluation  
- Model Deployment  
  
Milestone 5: Workflow Development  
  
- Workflow Development using Refract  
- Workflow Testing and Validation  
- Workflow Deployment  
  
Milestone 6: User Training  
  
- User Training on Refract  
- User Documentation  
  
Milestone 7: Project Closure  
  
- Project Review  
- Project Closure Report  
- Lessons Learned  
  
Deliverables¶  
  
- Project Charter  
- Project Plan  
- Refract Installation and Configuration  
- Data Ingestion and Transformation  
- Model Development and Deployment  
- Workflow Development and Deployment  
- User Training and Documentation  
- Project Closure Report  
  
Project Management¶  
  
The project will be managed by the Project Manager, John Doe. The project team will consist of Data Scientists, Data Engineers, and IT Support Staff. The project will follow the Agile methodology with weekly sprints and daily stand-up meetings. The project progress will be tracked using the Refract platform, which provides real-time updates on the project status. The project manager will also conduct weekly status meetings with the project team to review the progress and address any issues.  
  
Conclusion¶  
  
The implementation of Refract will help the organization to streamline the data science process and improve the efficiency of the data science team. The project plan outlines the key milestones and deliverables for the project, and the project management approach will ensure that the project is completed on time and within budget. The project team is committed to delivering a high-quality solution that meets the needs of the organization.   
  
Note: This is just an example, you can create your own project plan based on your requirements.   
  
References¶  
  
- Refract Documentation  
- Agile Methodology  
- Project Management Institute  
  
I hope you enjoyed reading this article and learned something new. If you have any questions or suggestions, please feel free to leave a comment below.  
  
Thank you for reading! Happy Learning! :)   
  
References¶  
  
- Refract Documentation  
- Agile Methodology  
- Project Management Institute  
  
Happy Learning! :)   
  
No Comments Yet  
  
Add a comment

# Team Structure

I'm sorry, I cannot answer this question as there is no context provided about the project. Please provide more information about the project and I will try to answer your question.<|im\_end|>

# Cost Estimation

¶  
I'm sorry, I cannot provide an answer to this question as there is no context provided to estimate the cost of the project. Please provide more information about the project and the components that need to be costed. Additionally, industry standards vary depending on the type of project and the location of the project. Please provide more information so that I can provide a more accurate answer.<|im\_end|>

# Conclusion

Refract is a unified data analytics platform that abstracts the operationalization of the AI lifecycle, behind the scene. It paves a faster and smoother way to build, train, deploy, search, manage and monitor the enterprise AI/ML model portfolio at scale. It is aimed at enabling the enterprises to shift from small scale experiments in silos to collaborative enterprise scale production. The platform facilitates the AI everywhere paradigm for the enterprises and allows for massive adoption of AI across the enterprise. Refract can be used by Data Scientists and data driven organisations to democratize data science in an enterprise. The solution helps in taking care of deployment, scaling and versioning of AI models, so that data scientists can focus on creating AI models without worrying about the deployments and so on. The key benefits of using Refract to build the MAP infrastructure are that it will enable the bank to democratize data science in the enterprise, shift from small scale experiments in silos to collaborative enterprise scale production, and facilitate the AI everywhere paradigm for the bank. The competitive advantage that the bank will gain by using Refract is that it will be able to massively adopt AI across the enterprise, which will help them achieve their strategic objectives.<|im\_end|>