A Report on

INPLANT TRAINING

AT

Jio Platforms Limited

Reliance Corporate Park, MUMBAI – 400701

Submitted By

Siddh Sanghvi

SAP ID: 70321019079

Under The Guidance Of

Prashasti Kanikar

Shri Vile Parle Kelvani Mandal's

Mukesh Patel School of Technology & Management Engineering

Department of Computer Engineering

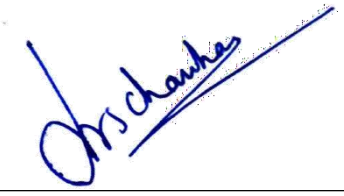
Vile Parle (w), Mumbai- 400056

Inplant Training Period

2nd January 2023 to 28th April 2023

NO OBJECTION CERTIFICATE

This is to certify that Siddh Sanghvi (SAP ID. 70321019079) student at Mukesh Patel School of Technology Management & Engineering, has successfully completed his/ her Semester **VIII** Inplant Training from 2nd January 2023 to 28th April 2023, at Jio Platforms Limited and has completed the tasks given to him during the Inplant Training. The Report is submitted in partial fulfillment of the requirement for B.Tech Integrated Computer Engineering. He has been allowed to include the relevant information from the Company for which we have no objection.



Signature & Stamp of The Industry

Name: Aditya Chauhan
Email ID: aditya3.chauhan@ril.com
Contact: +91 9326222811
Date: 02-05-2023

CERTIFICATE OF SUBMISSION

This is to certify that the report entitled 'Inplant Training at Jio Platforms Limited' is a bona fide work of Siddh Sanghvi (SAP ID. 70321019079) and submitted in semester **VIII** for the partial/complete fulfillment of B.Tech Integrated in Computer Engineering at, Mukesh Patel School of Technology Management & Engineering Mumbai during the academic year 2022-2023.

Signature of the Faculty Mentor	**Signature of the HoD**	**Signature of the Dean**

DECLARATION

I, Siddh Sanghvi SAP ID. 70321019079 a student of Semester **VIII** of B.Tech Integrated Program in Computer engineering humbly submit that I have completed from time to time the Inplant Training work as described in this report by my own skills and studied from 2nd January to 28th April. I have not copied the report or it's from any appreciable part from any other literature in contravention of academic's ethics. I declare that this written submission represents my ideas in my own words. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be the cause of disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Date: 28th April 2023

Siddh Sanghvi

Place: Mumbai

Name & Signature of the Student

- **Cover Page**
- **No Objection Certificate**
- **Certificate of Submission**
- **Declaration**

Chapter-1

Introduction:

Jio Platforms Limited is a significant Indian telecoms firm that offers a diverse variety of digital products and services. Jio Platforms was founded in 2010 by Mukesh Ambani, the Reliance Industries chairman, and is a division of one of the biggest private-sector corporations in India.

Jio Platforms Limited offers business-to-business (B2B) and business-to-consumer (B2C) solutions. A complete 5G solution is provided by the services, which also comprise a 5G radio, a full 5G core network, the AI/ML ATOM platform for 4G and 5G, the MANO cloud CNF orchestration tool, the ACI cloud infrastructure deployment tool, and cloud-native OSS platforms.

JPL has developed its own cloud-native probing solutions for radio and core networks in addition to these core platforms, which makes network debugging easier and eliminates the need for systems integration with probe suppliers.

Background of the Industry:

India's telecommunications sector has grown rapidly over the past ten years, with Jio Platforms setting the pace. Jio Platforms has been able to take advantage of India's sizable telecommunications industry by offering cutting-edge goods and services, and as a result, it has grown to become one of the nation's leading suppliers of digital solutions.

Services & Goods:

Jio, a wireless broadband service that offers high-speed internet, voice and video chatting, and digital content, is at the forefront of Jio Platforms' product lineup. The firm has introduced a number of digital apps and services in addition to Jio, including JioTV, JioCinema, JioSaavn, and JioMart. Customers can purchase digital material from these apps, such as music, films, and TV series. Jio Platforms' emphasis on innovative digital solutions equips it to satisfy changing demands.of its clients worldwide.

Fig 1. Jio Platforms Limited

Fig 2. Products & Services

About the Department

The department "New Initiatives AI/ML" is focused on exploring and developing cutting-edge technologies in the fields of Artificial Intelligence (AI) and Machine Learning (ML). The goal of the department is to leverage these technologies to create new products and services that can transform the lives of people and businesses in India. The department is staffed by a team of experienced professionals, who bring a wealth of knowledge and expertise in AI and ML to the table.

One of the key areas of focus for the department is computer vision. Computer vision is a field of AI that deals with the processing and analysis of images and videos. The department is working on developing new computer vision-based technologies that can help improve the accuracy of object recognition and tracking in various applications, such as autonomous monitoring, security systems, and retail.

In addition to these areas, the department is also exploring other emerging technologies such as Generative Adversarial Networks (GANs), Reinforcement Learning, and Deep Learning. The goal

is to stay at the forefront of AI and ML research and development, and to bring new and innovative products and services to market that can help transform the lives of people and businesses in India.

In conclusion, the New Initiatives AI/ML department at Jio Platforms is a key part of the company's strategy to continue its journey as a leader in the digital space. With a strong focus on cutting-edge technologies, the department is poised to play a critical role in the development of new and innovative products and services that can transform the lives of people and businesses in India.

Chapter-2

Technology/Equipment/Hardware

The New Initiatives AI/ML department at Jio Platforms uses a range of advanced technologies, equipment, and hardware to research, develop, and implement its AI and ML-based products and services. The key technologies used include :

Artificial Intelligence

Artificial intelligence (AI) is the emulation of human intelligence in devices intended to think and act like people. To create new goods and services, the division makes use of AI techniques like machine learning, deep learning, and computer vision.

ML, or machine learning

Machine learning is a branch of artificial intelligence that deals with creating algorithms that let computers learn from data.
To create new goods and services, the department makes use of ML approaches like reinforcement learning, unsupervised learning, and supervised learning.

Computer vision

The processing and analysis of images and videos falls under the umbrella of the AI discipline known as computer vision. One of the main technologies employed by the department is this one. The division makes use of methods like object recognition, image classification, and video analysis to develop new products and services.

Deep Learning (DL)

Deep learning is a subset of machine learning that uses multiple layers to progressively extract higher-level features from raw input. It is based on artificial neural networks with representation learning and can be supervised, semi-supervised or unsupervised.

VCS and IDEs

The department uses a variety of integrated development environments (IDEs), some of which are developer specific.

- Vim and Neo vim (Server),
- JetBrains PyCharm
- Nano (Server)
- Visual Studio Code and

The department's main version control system (VCS) is Gitlab.

Hardware

In terms of hardware, the department trains its AI and ML models on high-performance computing platforms such servers with graphics processing units (GPUs). In order to gather and analyse data for its computer vision-based goods and services, the department also makes use of specialised tools, such as computer vision cameras and sensors.

The department also has high-performance workstations with high-performance GPUs to use ML libraries that are GPU accelerated. Virtual access to cloud hardware and data is another usage for virtualization. Aggregated Virtual Control Platforms

Programming Languages/Tools/Software

The New Initiatives AI/ML department at Jio Platforms uses a range of programming languages, tools, and software to develop and implement its AI and ML-based products and services. Some of the key programming languages and tools used by the department include:

Python

Python is a high-level programming language that is commonly utilized in the fields of AI and machine learning. Python can be enhanced using the many open-source libraries that are readily available. It is one of the main languages used by the department for training and testing ML models, analyzing and cleaning data, visualizing data, etc. because of its simplified syntax, platform independence, and accessibility to large libraries.

Fig 3. Python

Fig 4 Conda

Scikit

Python has a machine learning library called Scikit-learn. SVMs, gradient boosting, k-means, random forests, and other regression, classification, and clustering methods are included in this system. A variety of tools and techniques are also available through DBSCAN for applications like classification, regression, clustering, and dimensionality reduction. A matplotlib-based Python data visualization library is called Seaborn. It offers a sophisticated drawing tool for creating eye-catching and educational statistical visuals.

Fig 5. Scikit Learn Library

Jupyter

Jupyter Notebook is an open-source web tool that allows the creation and sharing of documents that include live code, equations, visualisations, and narrative text. Block-by-block code execution and data visualisation are both performed using Jupyter Notebook kernels. To write, execute, and debug its programmes as well as to communicate its findings to others, the department uses Jupyter Notebook.

Fig 6. Jupyter Kernel

Seaborn

Seaborn is a Python data visualization library built on matplotlib. It offers a sophisticated drawing tool for creating eye-catching and educational statistical visuals.

Fig 7. Seaborn Library

Superset

Superset is quick, lightweight, intuitive, and rich with capabilities that make it simple for users of all skill levels to explore and visualise their data, from simple line charts to extremely comprehensive geographical charts.

Fig 8. Superset

PuTTY / Gitbash

A network file transfer tool, serial console, and free and open-source terminal emulator are PuTTY and GitBash. SCP, SSH, Telnet, rlogin, and raw socket connections are just a few of the network protocols that it supports.

Fig 9. PuTTY

Fig 10. Git bash

Vim / NeoVim

Vim is a highly customizable text editor designed to make writing and editing any sort of text incredibly efficient. The majority of UNIX systems and Apple OS X include it as "vi". To avoid the difficulty of constantly importing code from various platforms and then again debugging the code to verify whether it works on the server and to check if any problems have been created, Vim was used for server-side programming.

Fig 11. Vim IDE

Fig 12. NeoVim IDE

Docker

Docker is an open platform for developing, delivering, and operating applications. One can swiftly provide software by separating their apps from their infrastructure thanks to Docker. Development Environments can be run independently of libraries and other requirements thanks to containerization. The department uses Docker to containerize environments on distant servers for deployment, testing, and collaboration.

Fig 13. Docker

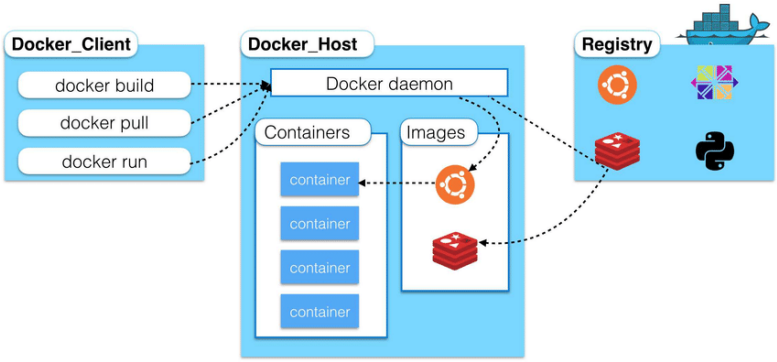


Fig 14. Docker Architecture

Flask

Flask is a web framework that is a Python module that allows you to easily construct web applications. Its core is compact and simple to extend; it's a microframework without an object relationship manager or similar capabilities.

Fig 15. Flask Framework

Pytorch

PyTorch is an open-source machine learning library that provides a seamless path from research to production. PyTorch may be integrated into a variety of ecosystems and is widely compatible. The division creates and employs its machine learning models using PyTorch.

Fig 16 PyTorch Framework

Other libraries and frameworks

The department uses a variety of other software and tools in addition to these programming languages and techn ologies, including data visualisation tools like Matplotlib and other frameworks like NumPy and Pandas.
To be at the forefront of AI and ML research and development, the department is always studying new and deve loping technologies such as Telnet, rlogin, and raw socket connection.

Fig 17. Various Python Libraries

Chapter-3

Projects Undertaken

The New Initiatives AI/ML department at Jio Platforms is involved in several projects to develop and implement AI and ML-based products and services. Some of the key projects assigned by the department include:

Machine Learning Approach to Attendance Score

Attendance is an important part of workplace since it allows employees to accomplish their work commitments while also maintaining a productive work environment. Although assessing and controlling employee timeliness can have a detrimental impact on productivity, expenses, and employee morale, it is a challenge faced by many organizations.

The main objective of this project is to develop a Machine Learning model that can accurately determine the attendance score of employees using only their entry and exit timings, without any prior knowledge of their work schedules or shifts. The model will be designed to classify users as either Fixed or Rotational, and then generate a score on their level of tardiness based on their attendance patterns.

Problem Definition

This project aims to develop a Machine Learning model that provides an attendance score for employees based solely on their entry and exit timings.

Constraints and Challenges

1. **Consecutiveness**

a. Punctual → If consecutively on time, then attendance score will be incremented but that increment will be at a steady rate as well as have a max cap amount which can be regenerated (optional)
b. Unpunctual → If consecutively tardy then attendance score will decrease at a linear rate.

2. Regularize

a. If considering stores or shops which have to be opened at a specific timing by a certain individual and if that particular individual is late then the attendance score of other individuals shouldn't be affected

3. Absent employees

a. when calculating Attendance score of a particular individual then need to consider absent employees and also should have an option to remove it from the number of company leaves allocated

4. shifts and field work.

a. In many cases different individuals have to step out of the office premises for company-related matters and based on that there attendance shouldn't be affected.

Tasks Undertaken

1. **Data Gathering:** Data Gathering involved the use of in-house data to create a dataset from which we were able to develop a more comprehensive understanding of employee shift patterns, behavior and performance.

a. In-house data was examined.

b. Relevant data was compiled into a dataset till the data requirement was satisfied.

2. **Data Preparation:** The dataset created was pre-processed was performed to ensure robustness by introducing a more diverse range of data to the model.

    a. Collected data was split into a train/val split structure (80%/20%)

    b. Data Cleaning techniques:

        i. Removal of Nan data

        ii. Removal of noisy data

        iii. Removal of outliers

        iv. Removal of duplicate data

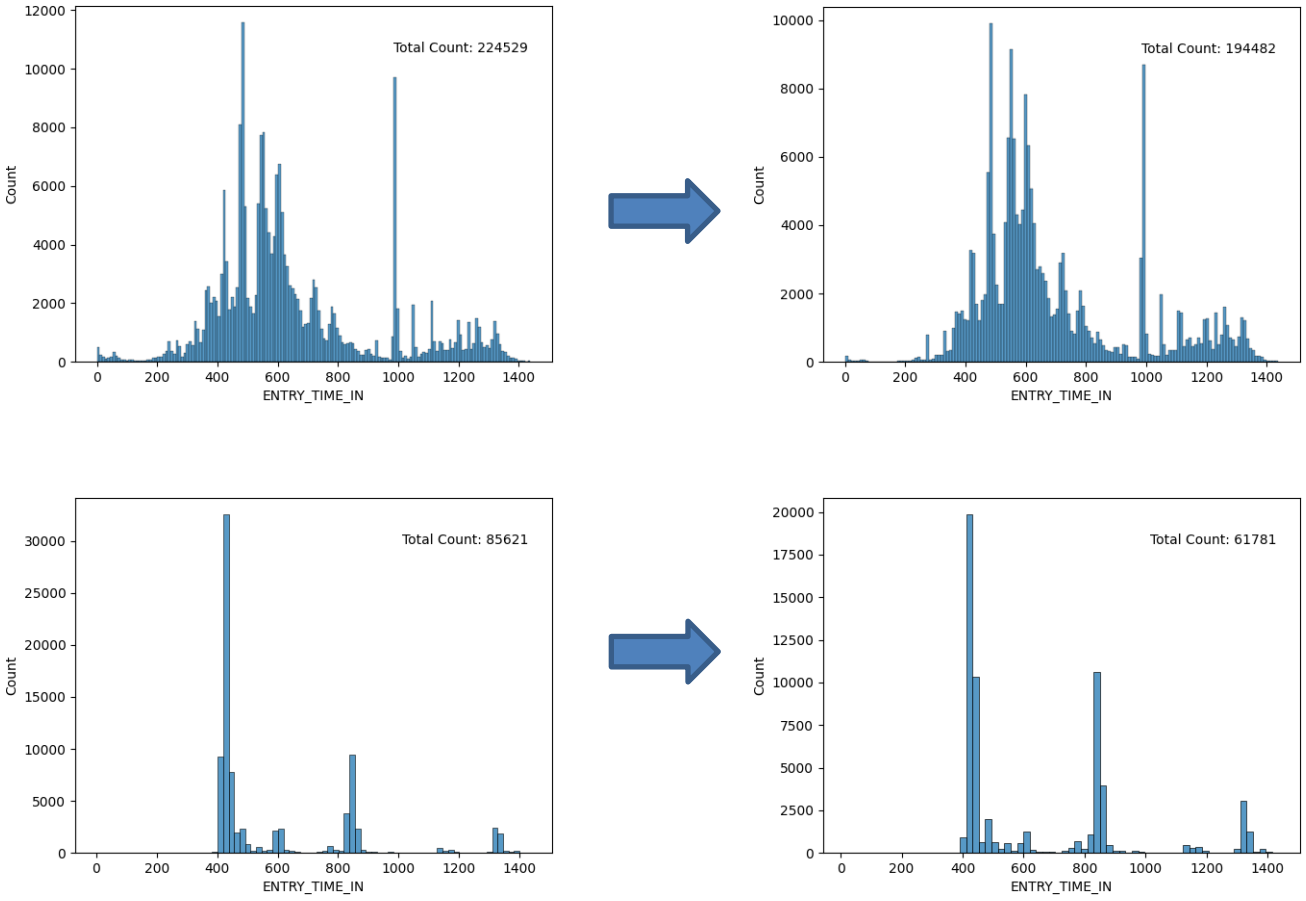


Fig 18. Count of Entries Before and After Data Cleaning

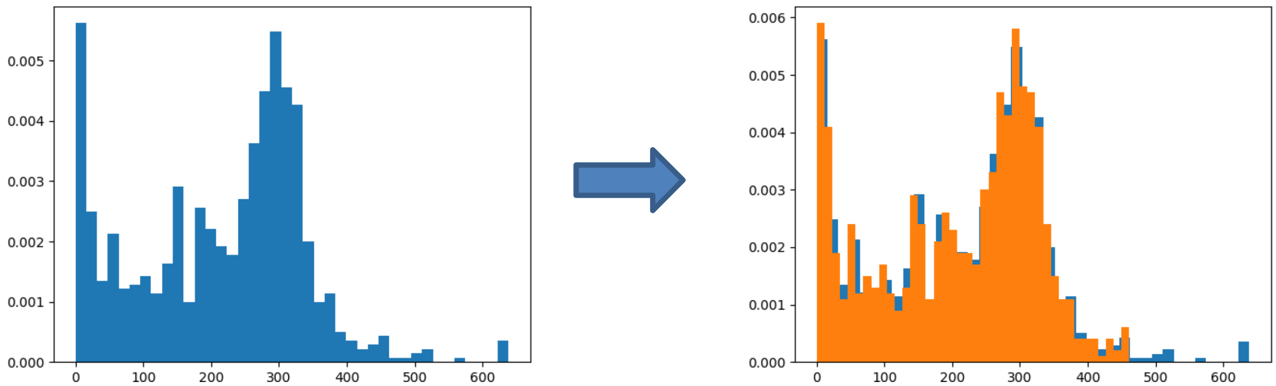


Fig 19. Pictorial Representation of data with and without Anomalises

c. Data transformation techniques:

i. Normalization

ii. One Hot encoding

iii. Feature Extraction

iv. Convert into NumPy array to pass in the model

0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.00
0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.00
0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.00
0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.00
0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.00
...	...	...	...	...	...	...	...	...	...	...	...	
0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.00
0.0	0.0	0.0	0.5	1.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.95
0.0	0.0	0.0	0.5	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.00
0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.00
0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.00

Fig 20. Tabular representation of ouput after applying Data transformation Techniques

3. **Model Selection:** Scikit-learn, a framework built over python was our primary Machine learning framework. We proceeded to test different clustering algorithms to find an optimal and efficient algorithm for classifying users:

a. K-Means Clustering

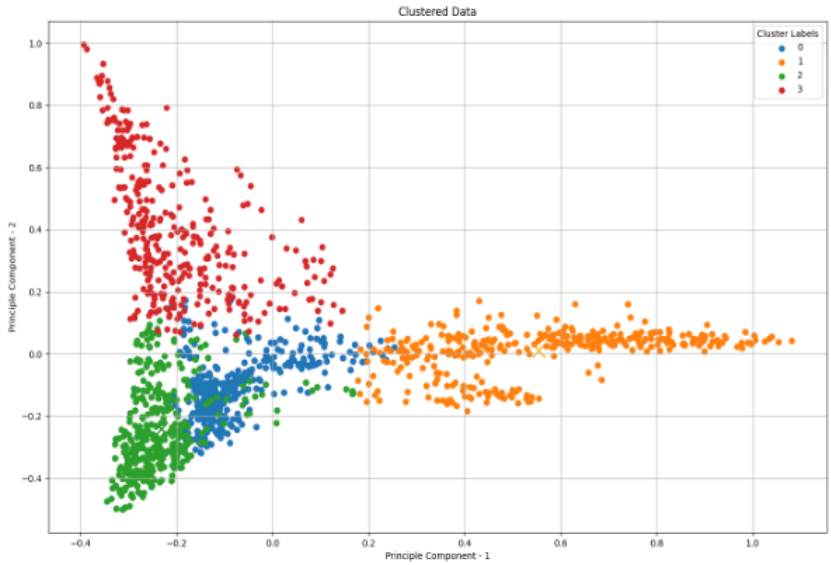


Fig 21 K- Means Clustering Model

b. K-Nearest Neighbor

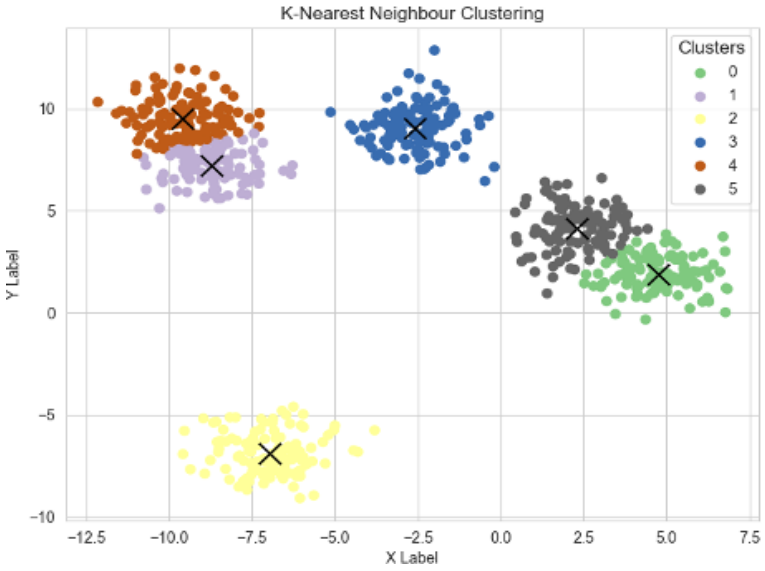


Fig 22. K-Nearest neighbour Model

c. Density Based Clustering (DBSCAN)

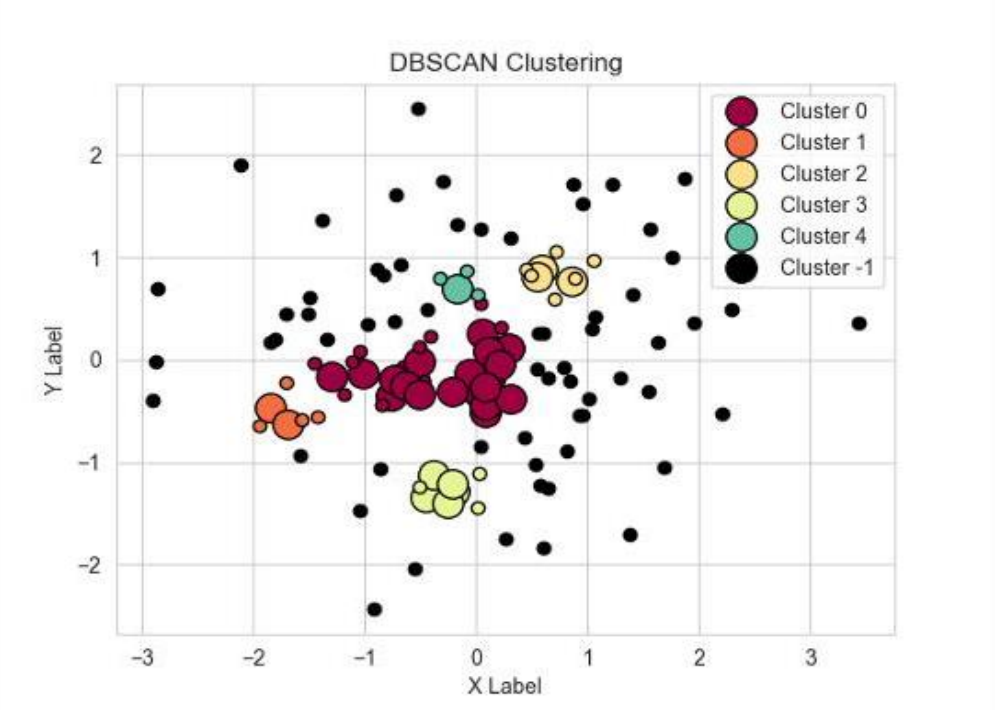


Fig 23. Density Based Clustering Model

d. Agglomerative hierarchal clustering

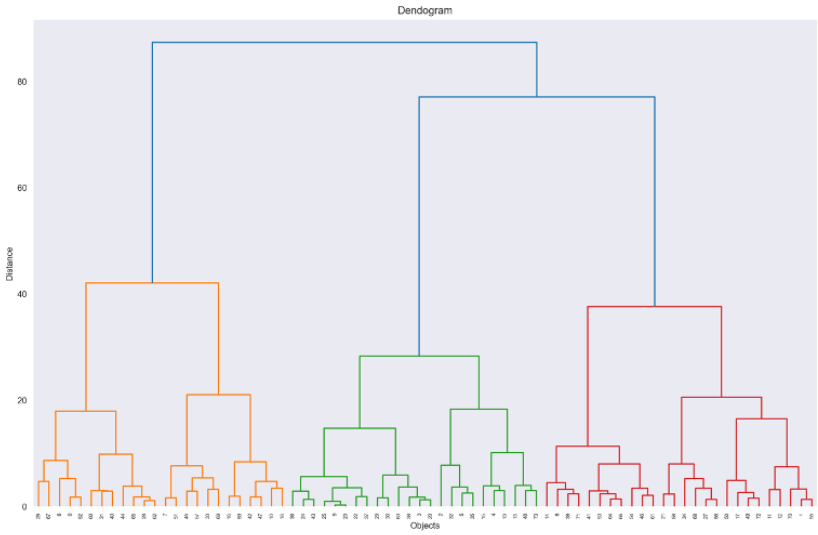


Fig 24. Agglomerative Hierarichal Clustering

4. **Model Training/Finetuning Pipeline:** Applied K-Means Clustering Algorithm to the dataset and parameters were tuned to obtain a model checkpoint that performs as per our requirement.

    a. Different distance calculation metrics were used such as Euclidean, Manhattan and Cosine

    b. Loaded the model on a preset range of 1 – 10 number of clusters and chose the best output on different datasets

    c. Best performing tuned models were selected for further fine tuning over in-house dataset.

5. **Model Evaluation:** : To evaluate the performance of the model, we utilized both the silhouette score and the elbow method. These metrics provided insight into the quality of the clustering and helped determine the optimal number of clusters for accurately identifying shift information.

    a. Model was tested on test dataset

    b. Elbow Method is a graphical representation of finding the optimal 'K' in a K-means clustering.

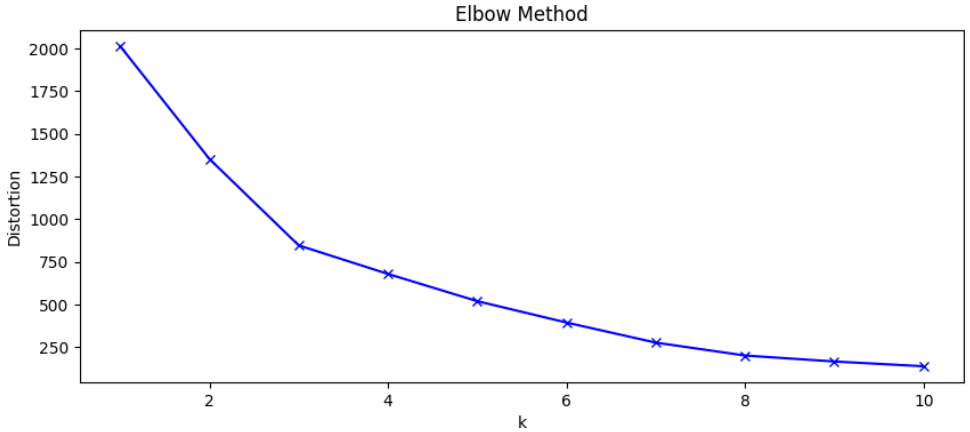


Fig 25. Elbow method for finding optimal number of clusters

c. silhouette score was used to measure of how similar an object is to its own cluster (cohesion) compared to other clusters (separation).

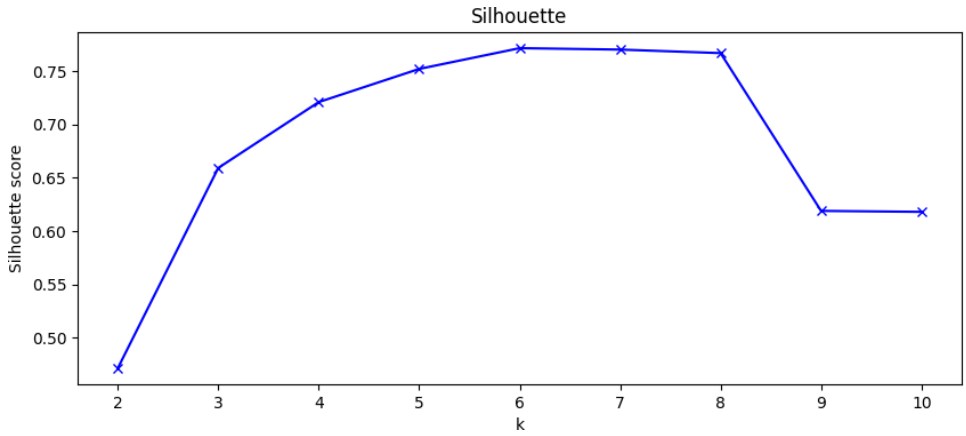


Fig 26. Silhoutte Score for finding optimal number of clusters

d. Used manual inference to extract meaningful insights from the model, which allowed for a deeper understanding of the data and helped fine-tune the model's parameters for optimal performance.

6. **Model Testing**: During the Model Testing phase, we classified each data point into a cluster and subsequently examined the clusters to ensure accurate classification of all data points.

a. Created scripts to annotate the data and then manually compared the annotations with the model's predictions

	NAME	Prediction	Cluster
0	AMLESH KUMAR YADAV	2	2
1	ANKUSH	3	3
2	ASHOK SHAIKH	1	0
3	Abhijit kamble	2	2
4	Abhishek Mishra	2	2
...	...	...	...
263	shiv chopal	0	0
264	smiti aawahad	1	0
265	surjeet kumar	1	1
266	vikram singh	1	1
267	yog raj	1	1

Fig 27. Predictions made by Clustering Model

b. Used Pairwise distance metric to evaluate each data points distance to a cluster and assigned the cluster with lowest distance

7. **Model Deployment**: We developed a custom Flask server and modularized the code to deploy the model in pickle format. We also created an API to convert the model into a SaaS solution. To test the API's functionality, we designed and executed a series of integration tests using a simulated client, which emulated real-world user requests. These tests helped ensure that the API was working correctly and provided reliable results..

Fig 28. Screenshot of Flask app deployed on server

Deep Learning Approach to Attendance Score:

The aim of this project is to build a deep learning model using a fully connected CNN that can classify users as either having fixed or rotating shifts. This binary classifier will enable the calculation of attendance scores based on entry and exit timings, providing organizations with an effective tool for workforce management.

Problem Definition

The problem was defined as the need for a more accurate system that can recognize fixed or rotational user.

Tasks Undertaken

1. **Data Collection:** Inhouse dataset, which was created for machine learning approach was used to develop the classification model.
    a. Data from various sites was examined.
    b. Relevant data was compiled into a dataset till the data requirement was satisfied.
2. **Data Preparation:** The dataset created was pre-processed to ensure robustness by introducing a more diverse range of data to the model.
    a. Collected data was split into a train/test split structure (80%/20%)

b. Implemented a python GUI based scripts which would display a plot of each users normalized data and then we used buttons to append whether user is fixed or rotational to the dataset

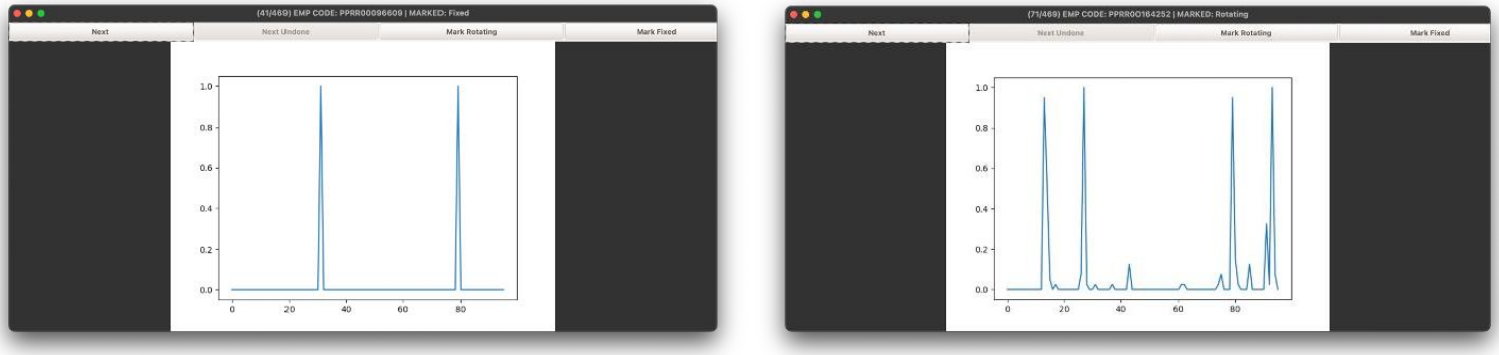


Fig 29. GUI interface of application developed to dynamically annonate data

c. Data was labelled as Fixed and Rotational data.

d. Data processing techniques were applied such as:

   i. Data Normalization

   ii. One Hot Encoding

3. **Model Selection:** PyTorch was our primary deep learning framework. To identify the optimal model architecture, we conducted extensive experimentation, varying parameters such as the number of layers, the number of neurons in each layer, and the activation functions.

   a. Research regarding current best performing text based classifier models was conducted.

   b. batch size, learning rate, and dropout rate was also considered to ensure that the model was accurately trained on the underlying data.

4. **Hyperparameter Tuning:** After training the initial model, we fine-tuned it by adjusting the hyperparameters to optimize its performance

    a. used techniques such as grid search or random search to find the best combination of hyperparameters

    b. The hyperparameters (optimizer, learn rate, loss function) of the best checkpoints were tuned to further improve model performance.

    c. Best performing tuned models were selected for further fine tuning over in-house dataset.

(Further plans)

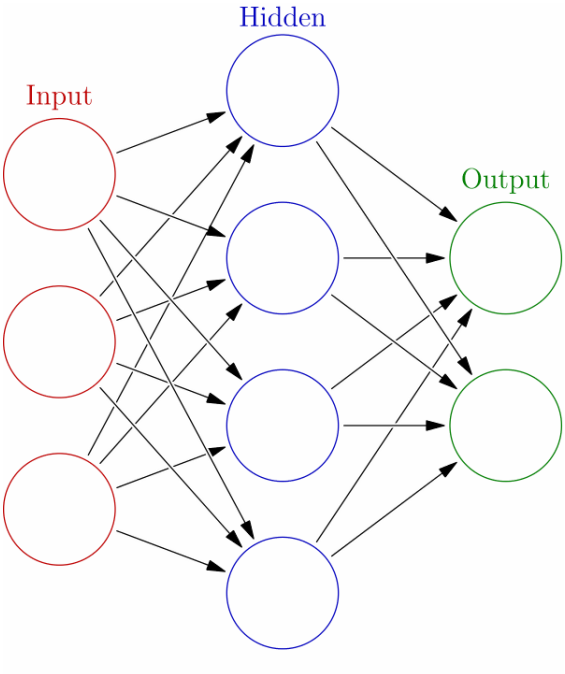


Fig 30. Binary Classifier Architecture

Ongoing development: As part of ongoing development, I plan to continue improving the model's performance by experimenting with different architectures, pre-processing techniques, and hyperparameters. I also plan to incorporate additional features such as user demographics and job roles to further improve the accuracy of the model. Additionally, I will continue to monitor the model's performance and update it as needed to ensure that it continues to provide accurate predictions.

Research Projects:

Website Development Deck

This initial project involves conducting research and understanding the pipeline and flow of the website development process in order to identify what processes in the flow can be automated or assisted using AI.

Such automation and assistance can help improve developer productivity and efficiency as well as monitor and improve code and product quality. A presentation was created and was presented to our department superiors.

Software Factory Implementation

This project involves the development of a software factory that is designed to streamline the development and deployment of AI and ML-based products and services. The software factory is based on the Scaled Agile Framework (SAFe) and provides a standardized approach to software development, enabling the department to deliver high-quality products and services in a timely and efficient manner.

Thorough research was conducted in order to establish the architecture, flow and components of the software factory and a presentation of the topic was created and presented to our superior of the department.

Scaled Agile Framework (SAFe)

SAFe is a framework for scaling Agile development practices across multiple teams. The department has adopted the SAFe framework to ensure that its projects are delivered on time and with high quality, while also ensuring that teams are aligned and working together effectively.

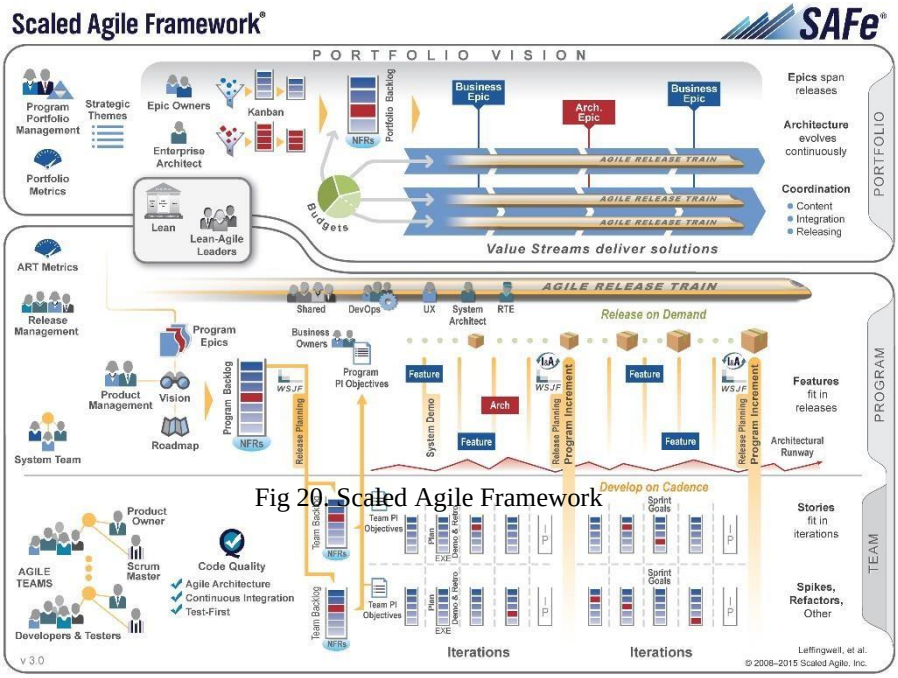


Fig 20. Scaled Agile Framework

Miscellaneous Projects:

Remote Sensor Onboarding

Sensor onboarding refers to the process of making installed sensors visible and accessible via inhouse software by use of specification configurations such as:

1. Admin Credentials
2. IP addresses

Once onboarded sensor data is accessible and recorded into the inhouse software.

Remote Sensor Certification

Sensor certification refers to the process of verifying the installation and functioning of the sensors and ensuring the sensors are performing as specified. Using our inhouse software interface sensor data was examined and sensors showing faults/anomalies/no recording were flagged. Criteria for unusual data was defined and faulty sensors were categorized as per the defined criteria.

Chapter-4

Learnings in the Inplant Training

As an intern in the New Initiatives AI/ML department at Jio Platforms, I have had the opportunity to gain a wide range of technical skills and knowledge. Some of the key technical learnings that I experienced include:

Agile development methodologies

By working in an Agile development environment and following the Scaled Agile Framework (SAFe), I am learning how to work in a fast-paced, iterative development process. I also gained an understanding of how to work in a cross-functional team and how to deliver high-quality software on time.

AI and ML technologies

I have the opportunity to work with a range of AI and ML technologies, including deep learning and computer vision.

I am learning how to build and train models, how to evaluate their performance, and how to integrate them into real-world applications.

Software development best practices

I am learning about software development best practices, including code quality, testing, and continuous integration and delivery.

I am also gaining an understanding of how to work with source control systems such as Git, and how to use tools such as JioWorks and Kanban boards to manage software development projects.

Modular code generation

I am learning about modularizing my codes so that they can be reused in different applications and processes.

I have understood best coding practices and syntax and also understood how to write more developer friendly code using OOPs.

Data science and engineering

I am gaining an understanding of how to work with large datasets, how to clean and preprocess data, and how to perform exploratory data analysis. I am also learning how to use data visualization tools such as Matplotlib and Seaborn to visualize and understand your data.

Problem-solving skills

I have the opportunity to work on real-world projects, and as a result, I am developing problem-solving skills. I am learning how to identify and solve complex technical problems, and how to communicate my solutions effectively to stakeholders.

By working on these projects, I gained a solid foundation in AI and ML development and software engineering, which would be valuable for your future career. Additionally, I have the opportunity to network with experienced professionals in the field, which could lead to future opportunities for growth and advancement.

Chapter-5

Conclusion and Future Plan of Action

As an intern in the New Initiatives AI/ML department at Jio Platforms, I have the opportunity to work on exciting projects and gain valuable experience in the field of AI and ML. In conclusion, I would likely have the following takeaways and future plans of action:

Valuable technical skills and knowledge

I have gained a solid foundation in AI and ML development, software engineering, and data science, which would be valuable for your future career.

Networking opportunities

I have had the opportunity to network with experienced professionals in the field, which could lead to future opportunities for growth and advancement.

Passion for AI and ML

My experience in the New Initiatives AI/ML department has sparked and further solidified my passion for AI and ML, and I have a desire to continue learning and exploring these exciting fields.

Based on these experiences, my future plan of action includes:

Building my portfolio

I would want to continue building my portfolio of AI and ML projects, both through my work at Jio Platforms and through personal projects.

Networking with professionals in the field

I would want to continue building my network of professionals in the field, through online communities, professional organizations, and conferences.

Staying up-to-date with the latest developments

I would want to stay up-to-date with the latest developments in AI and ML, through reading

industry publications, attending conferences, and participating in online communities.

In conclusion, my experience as an intern in the New Initiatives AI/ML department at Jio Platforms is a valuable steppingstone for my future career, and my future plans of action would likely reflect your passion for AI and ML and my desire to continue learning and growing in this exciting field.

CERTIFICATE OF INPLANT TRAINING

This is to certify that Mr. Siddh Sanghvi SAP ID. 70321019079 of Semester **VIII** of B.Tech Integrated Program Computer Engineering has attended the Inplant Training for 83 days out of 83 days at Jio Platforms Limited during the Period: from 2nd January 2023 to 28th April 2023.

Signature of the Student

Signature of the Industry Mentor
Name: Aditya Chauhan
Email ID: aditya3.chauhan@ril.com
Contact: +91 9326222811
Date: 02-05-

Forwarded with compliments to the HoD for necessary action.

The details of working at organization:

Present: 83 Days

Leave: 0 Days

Absent: 0 Days

Total: 83 Days