REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

REVA Academy for Corporate Excellence (RACE)

Key Driver Analysis for Customer Decommissioning using Machine Learning

ARUNA K

SRN: R19MBA54

Date: 20.08.2022

PGDM/MBA in Business Analytics

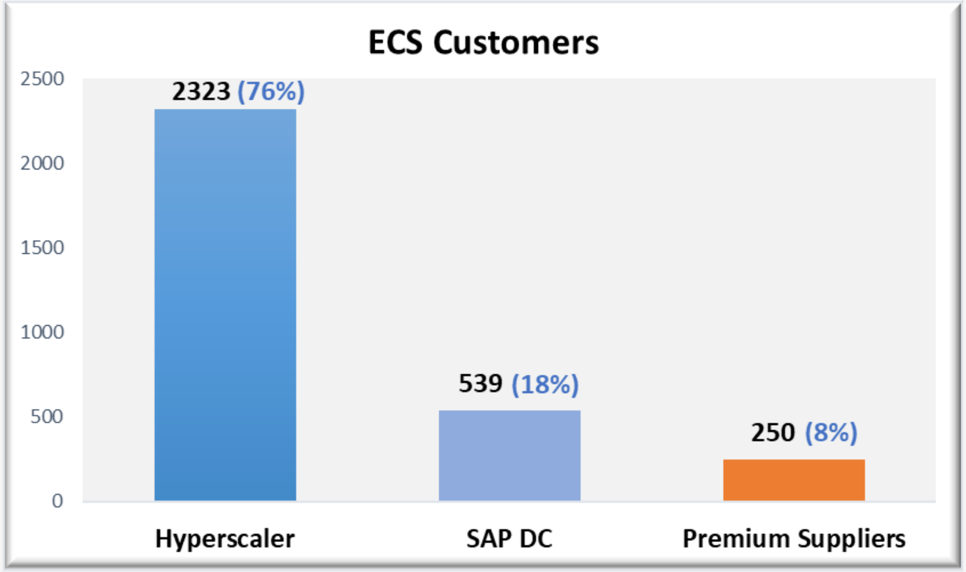
Capstone Project Presentation
Year: I

race.reva.edu.in

Agenda

SAP - ECS

- SAP Enterprise Cloud Services (ECS) delivers all private cloud ERP solutions to customers.
- ECS serves **3,053** customers with an increase of **3%** MOM.
- SAP operates on close to 30 data centers and hyperscalers including cloud service providers such as **AWS, Azure, and GCP**.



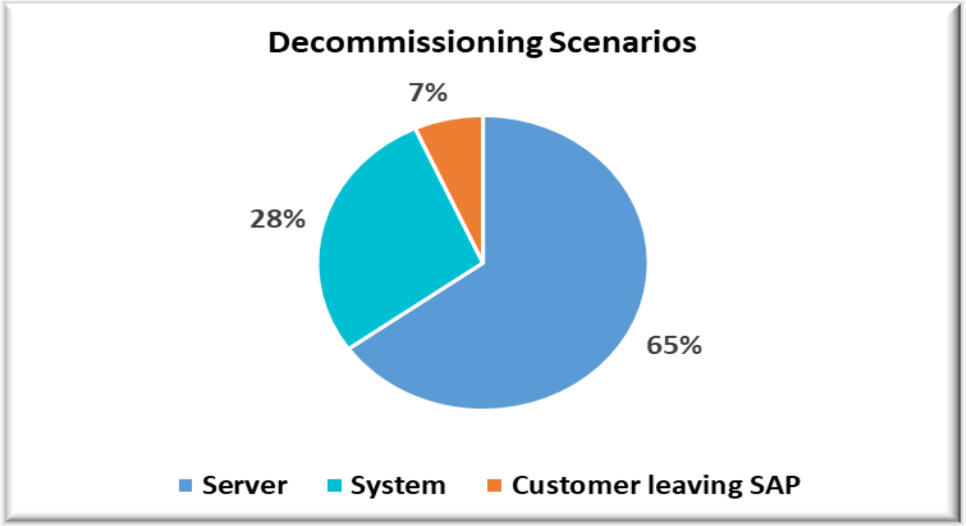
ECS - Decommissioning

- Decommissioning is the last phase in the life cycle of a customer system and is the permanent removal of a configuration item from the live customer environment.

- ECS performs customer solution decommissioning upon the contract end date of the customer.

- Decommissioning process provides a structured way to decommission a system or a server or network.

- 7% churn rate is low compared to industry standards indicating that contract renewals are key to the business.



Why decommissioning?

- Decommissioning of customer systems upon contract end date saves hardware and infrastructure cost for ECS.

State of Art

AUTHORS, PAPER PUBLISHED, YEAR OF PUBLICATION	RESEARCH WORK
SIMS Lifecycle Services – "Solutions for Global Data Center Decommissioning", 2020	Research has been done on all pivotal areas in the global data center decommissioning process, however the focal point is only on data center decommissioning.
SIMS Lifecycle Services – "Sustainable Data Center Decommissioning", 2020	With the growing data needs, as well as hyperscalers and data centers in existence, research on ensuring the data centers are environmentally sustainable is the main goal of this research.
Adnan Masood & Adnan Hashmi – "AIOps: Predictive Analytics & Machine Learning in Operations", 2019	The paper addresses how to handle more incidents with shorter service-level agreements (SLAs), respond to these incidents more quickly, and improve on key metrics, such as mean time to detect (MTTD), mean time to failure (MTTF), mean time between failures (MTBF), and mean time to repair (MTTR).
Owusu Nyarko-Boateng et.al – "Using machine learning techniques to predict the cost of repairing hard failures in underground fiber optics networks", 2015	The paper investigated the cost of repairing underground fiber cable failures and then used feedforward neural networks (FFNN) and linear regression to predict the cost of repairing future faults. The result of the model predicts the costs of repairing underground optical networks before the fault occurs.

- A detailed study on best practices for the general decommissioning process and reduction of turnaround time have been covered as part of the literature review.
- Machine learning is used in the decommissioning process for the reduction of Mean Time To Resolve (MTTR) which is a hybrid approach and novelty of this project.

Problem Statement

2X delay in time taken to perform end-to-end decommissioning process affecting the business of ECS.

Loss of ~150K Euros/Month has been incurred by ECS for retaining the customer systems even after the contract has been ended.

There is no ticketing dashboard to track the status of decommissioning.

REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956, Approved by AICTE, New Delhi

Project Objectives

Three major objectives of this study are as follows:

- Analyzing the Mean Time To Resolve (MTTR) metric to identify the delays using decision trees and KNN Classifier

Key Drivers

- Savings up to ~150K Euros/Month by optimizing the process thereby reducing delays

Cost Savings

- Develop a JIRA ticketing dashboard for the ECS leadership team collating the ticketing KPIs

Dashboard

REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

Project Methodology

CRoss Industry Standard Process for Data Mining Framework

Business Understanding

Customer systems are decommissioned once the contract is terminated. Delays in decommissioning process impacts cost.

Data Understanding

The analysis is based on the JIRA ticketing dataset consisting of 1624 tickets created for decommissioning requests.

Data Preparation

The dataset has been split into training and test data containing MTTR, a continuous variable indicating the days taken to resolve the ticket.

Modeling

The decision Tree algorithm has been considered to predict the key drivers and their relationship with MTTR. KNN classifier predicts the MTTR with the highest accuracy.

Evaluation

Evaluation of the model results and review the steps executed to be certain that it properly achieves the business objectives.

Deployment

The outcome of the model shall be used to forecast delays for the future data with the recommendations provided to the relevant teams to optimize the decommissioning process.

Business Objectives

Business Model Canvas

Key Partners
- ✓ Sub-contractors
- ✓ GCS team for Network & Backup deletion
- ✓ Server Management
- ✓ Database Deletion Team
- ✓ Other Support Teams

Key Activities
- ✓ Data Export
- ✓ Decommissioning

Key Resources
- ✓ Decom Team
- ✓ CDM/TSM
- ✓ Other support teams

Value Proposition
- ✓ On-time decommissioning
- ✓ Reduce backlogs
- ✓ Revenue savings
- ✓ Process Optimization
- ✓ Track metrics and KPIs

Customer Relationships
- ✓ Eliminate wrong decommissioning
- ✓ Reducing delays

Channels
- ✓ CDM/TSM
- ✓ Service Requests

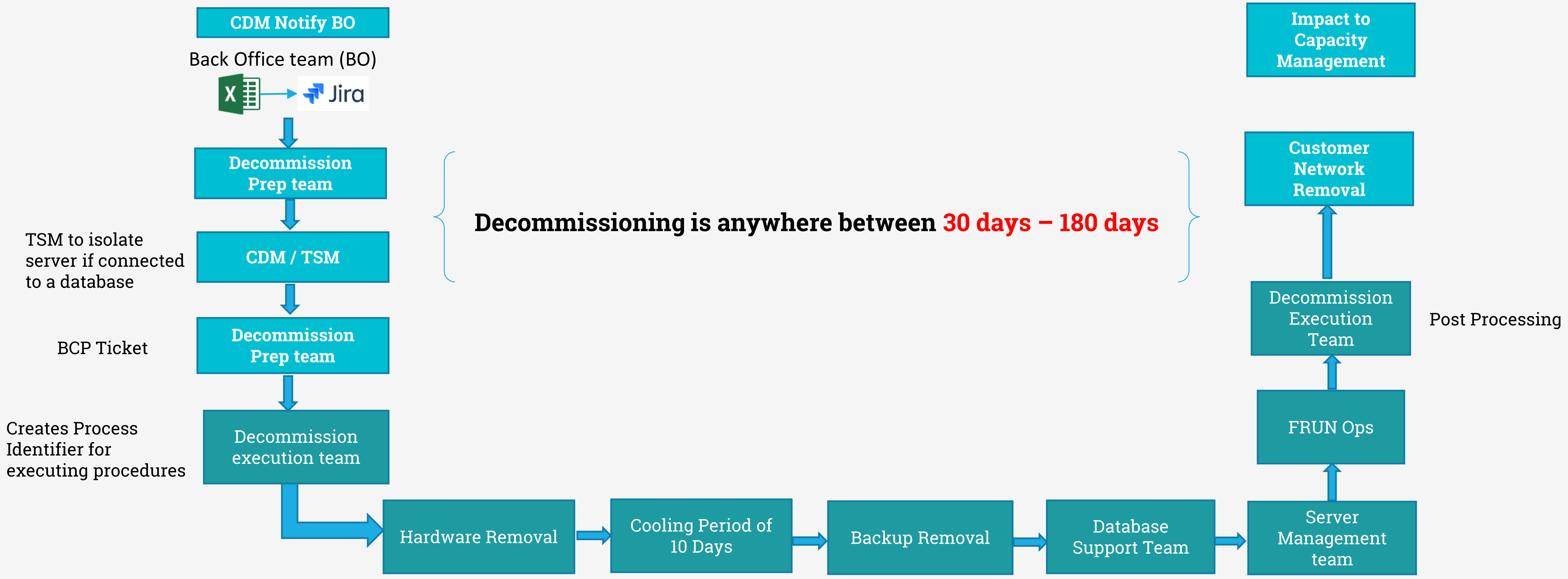
Customer Segments
- ✓ Customers on SAP landscape
- ✓ Customers on Hyperscalers (AWS, Azure, GCP)

Cost Structure
- ✓ Sub-contractors
- ✓ GCS team for Network & Backup deletion
- ✓ Salaries to Employees

Revenue Streams
- ✓ Licensing fee
- ✓ Subscription fee
- ✓ Hybrid Pricing

REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

End to End Decommission Process



REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956, Approved by AICTE, New Delhi

Critical features for MTTR Analysis

Data Dictionary – Features considered in the dataset

Features	Description
Key	Unique tickets created in JIRA
Summary	Ticket description
Status	Ticket status
Assignee	Last assignee worked on the ticket
Created	Ticket creation date
Resolved	Ticket resolution date
Updated	Ticket last updated date
Priority	Priority of the ticket
Sub-Contractor	Vendor working on the ticket
Types of Decommission	Server, System, Network or customer leaving SAP
No. of Servers	Total number of servers to be decommissioned
Cost (Euros)	Cost for maintaining the server
MTTR (Days)	Mean Time to Resolve (Resolved – Created)
Isolation (Days)	Number of days taken to isolate a server
Downsizing (Days)	Number of days taken to downsize the server
Cooling Period (Days)	Number of days the server was kept on cooling period
Backup Removal (Days)	Number of days taken to delete data backup
Database Removal (Days)	Number of days taken to delete database
Volume's deletion (Days)	Number of days taken to delete server volume
FRUN (Days)	Number of days taken to delete FRUN
Network (Days)	Number of days taken to delete Network

Key independent variables considered are –

- Sub-Contractor
- Cost in Euros 5K per TB per Month
- E2E Decommission Process in Days

Data Preparation

Data Type	Time window considered	Number of tickets (All)
JIRA Tickets	January 2021 – July 2022	1634

Data Preparation Steps:

- Dataset is split into two categories training, and test set.
- The training and test dataset contains MTTR, a continuous variable providing the days taken to resolve the ticket and complete decommissioning.
- The test set is used as the deployment data for the re-validation of the model.
- Key variables considered are continuous variables. Sub-Contractor is a categorical variable.

Reasons for dropping features

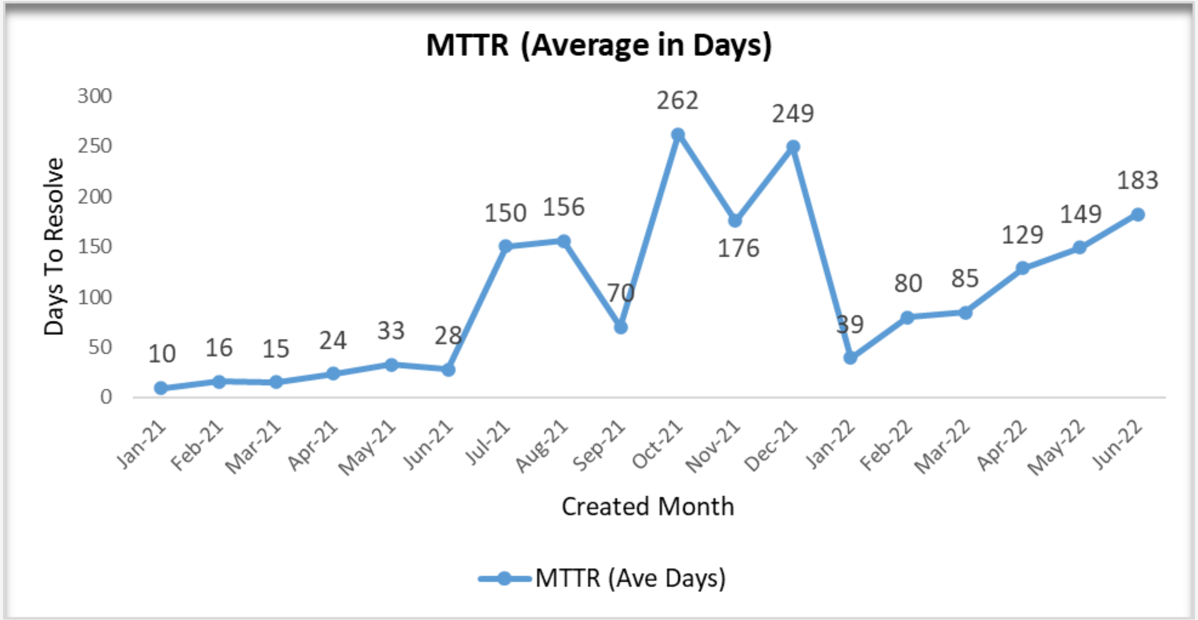
Features	Comments for dropping the features
Key	Unique tickets for decommissioning requests
Summary	Dropped as it's descriptive field
Assignee	Dropped as not relevant
Component/s	Dropped as region is not a factor
Priority	Post factor feature, hence dropped
Number of Systems	Dropped as total number of servers has been retained
Number of Network	Insignificant as customer leaving SAP is marginal

REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956, Approved by AICTE, New Delhi

Cost and Time - Delay Analysis

Servers yet to be decommissioned and the cost associated

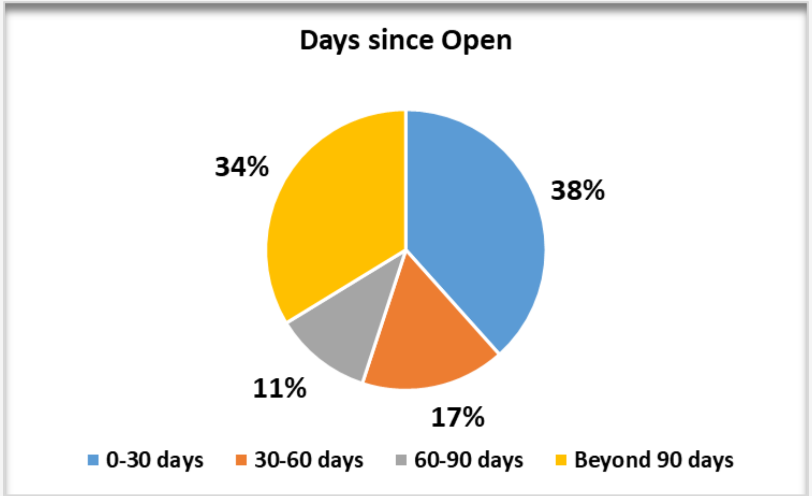
Year	No of Servers	Cost (Euros)
2021	255	€ 12,75,000.00
2022	900	€ 45,00,000.00
Grand Total	1155	€ 57,75,000.00



Increasing trend of MTTR indicating delays in the decommissioning process.

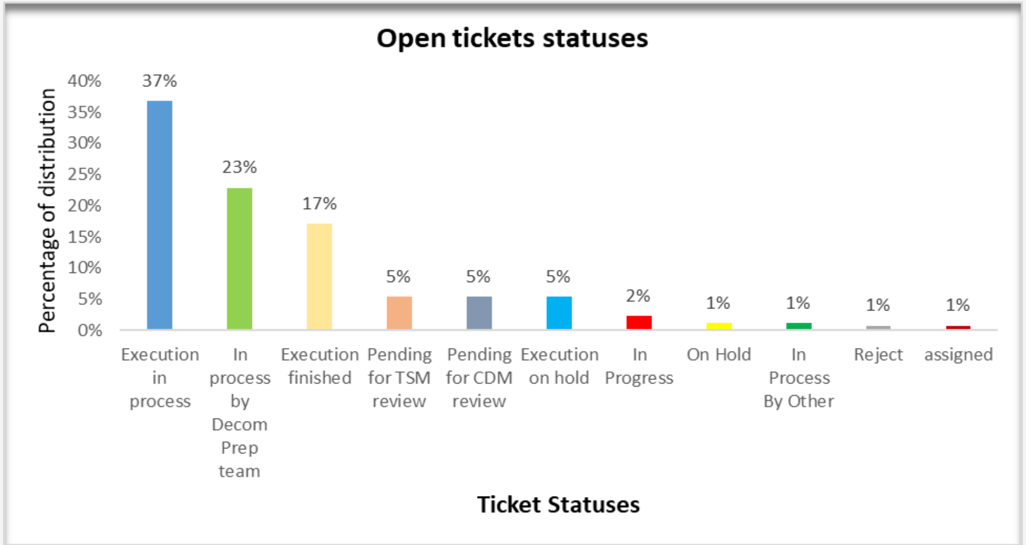
Open Tickets Analysis

Tickets Open (in Days) since created



Over **60%** of tickets open beyond 30 days are considered backlog

Tickets in various Statuses (Open)

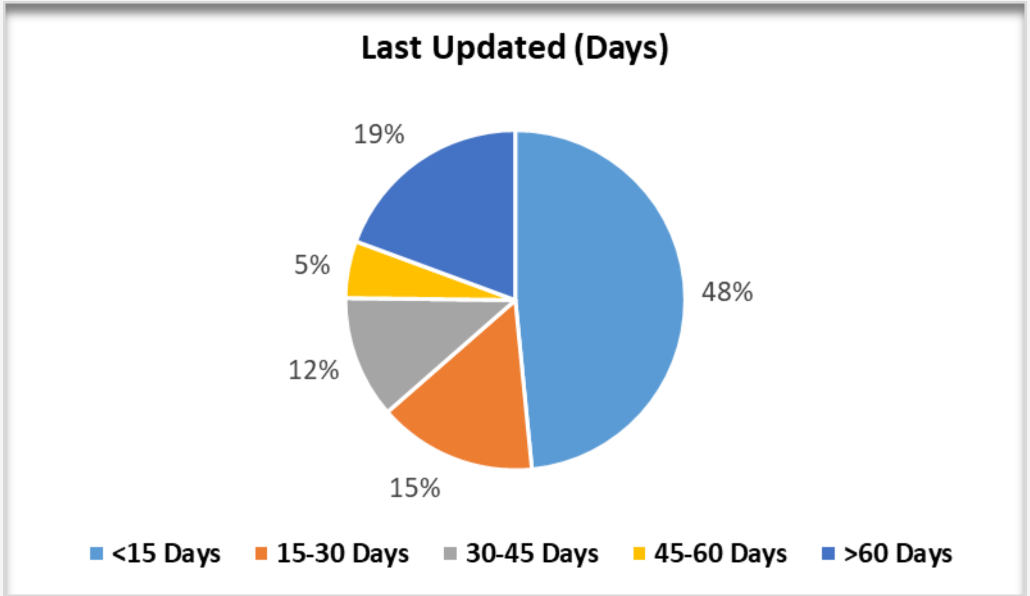


- **60%** of "Open" tickets are under the "Execution in process" status.
- **15%** of tickets need immediate attention as the tickets are awaiting responses from role owners (CDM, TSM).

Tickets Updated & Vendor Analysis

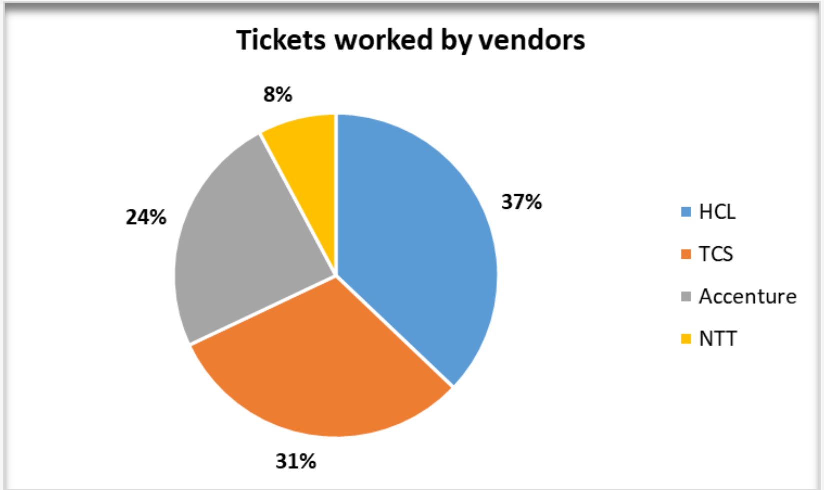
Established as per the section 2(f) of the UGC Act, 1956, Approved by AICTE, New Delhi

Days since tickets last updated



A total of **67%** of tickets that are not updated for 30 days need immediate attention from the team.

Tickets Open – Vendor (Sub-Contractor) wise

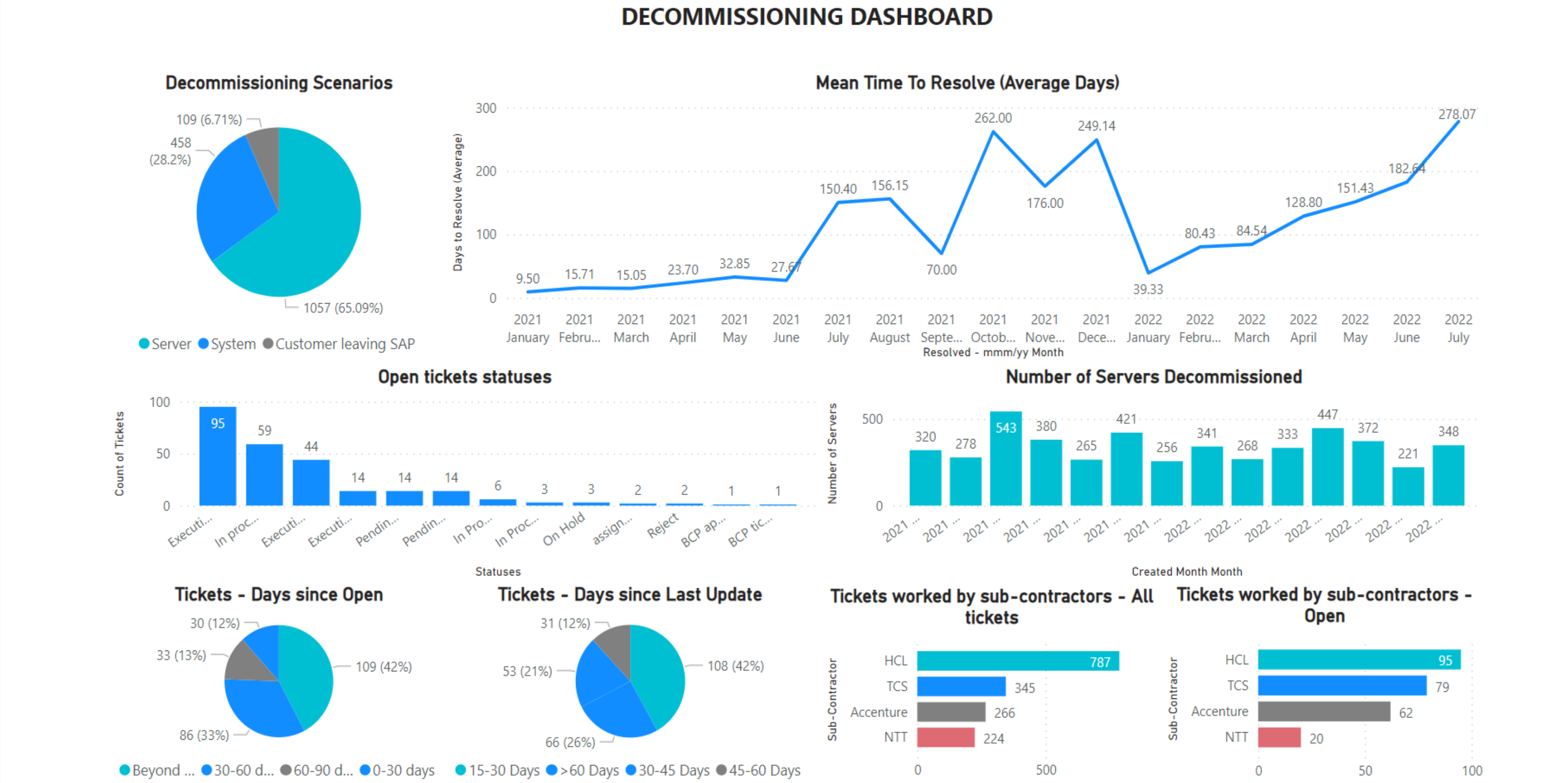


- Shows the distribution of tickets worked by Sub-Contractors.
- It is imperative to study if the delays are caused by Sub-Contractors.

Decommissioning Ticketing Dashboard

REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956, Approved by AICTE, New Delhi

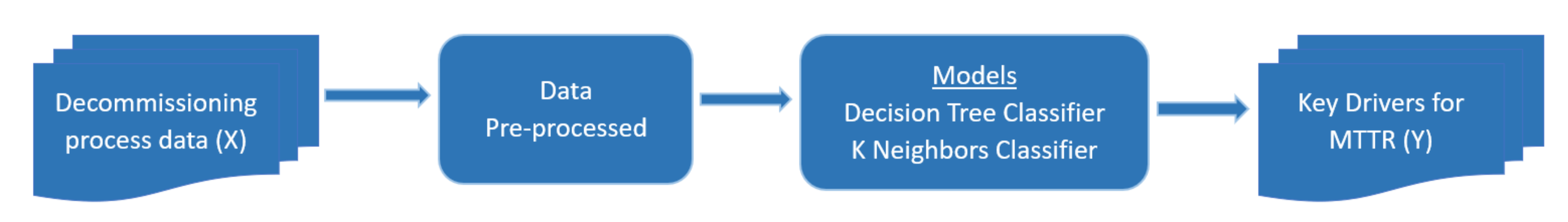
Power BI Dashboard - Decommissioning KPIs



Link to the dashboard: https://app.powerbi.com/groups/me/reports/07db46f7-7110-4768-bced-a659fdc73ee8/ReportSection06e95c4d41cc0b60ce48

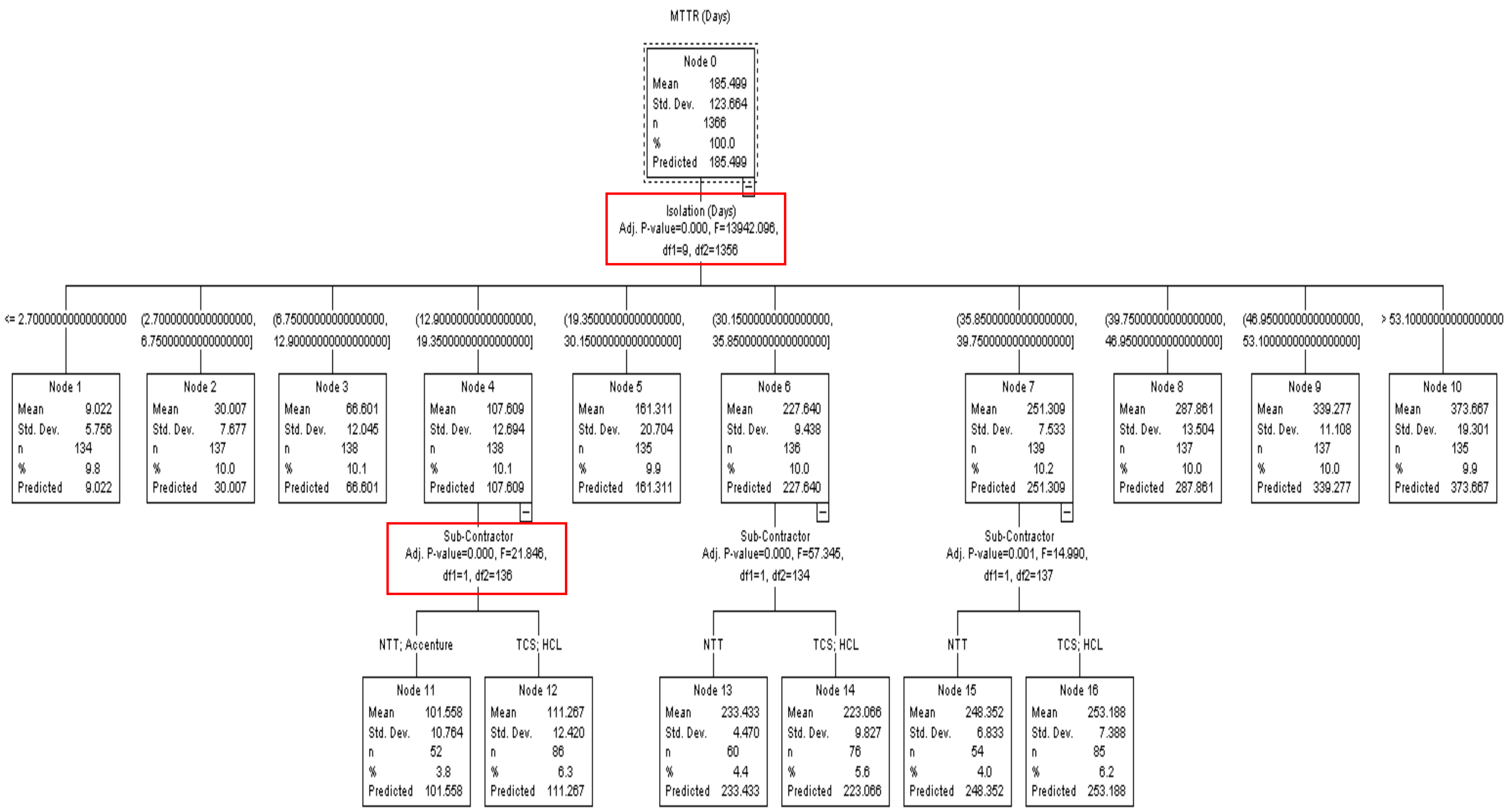
Modeling

Overview of the dataflow into the Machine Learning Data Model



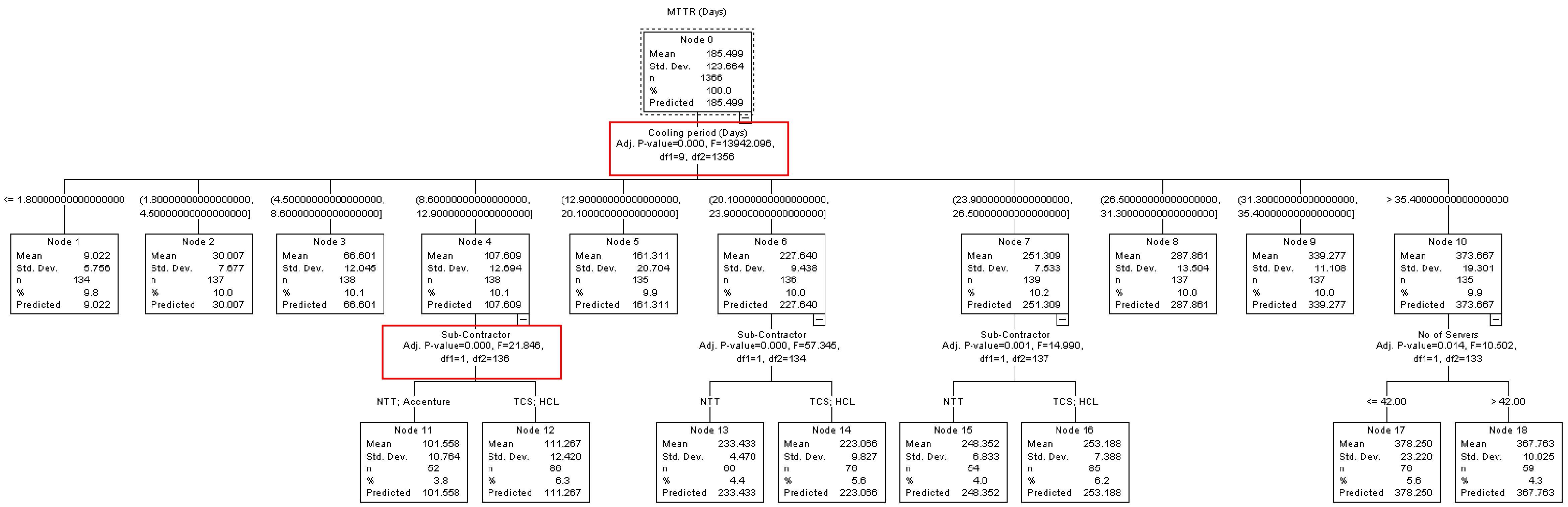
- Based on the problem statement and the data availability, the Decision Tree algorithm has been considered to predict the key drivers and their relationship with MTTR.
- KNN Classifier has been considered to forecast the MTTR as an alert mechanism for new tickets.
- 70% of data is used as training sets and 30% of data is used as testing sets.
- The training set is used to train the decision tree model with certain features of tickets, such as the number of servers, server isolation, virtual machine downsizing, cooling period, backup removal, database removal, volume deletion, FRUN, and network deletion.
- A Testing set is used to evaluate the performance of classifiers.

Decision Tree Output 1



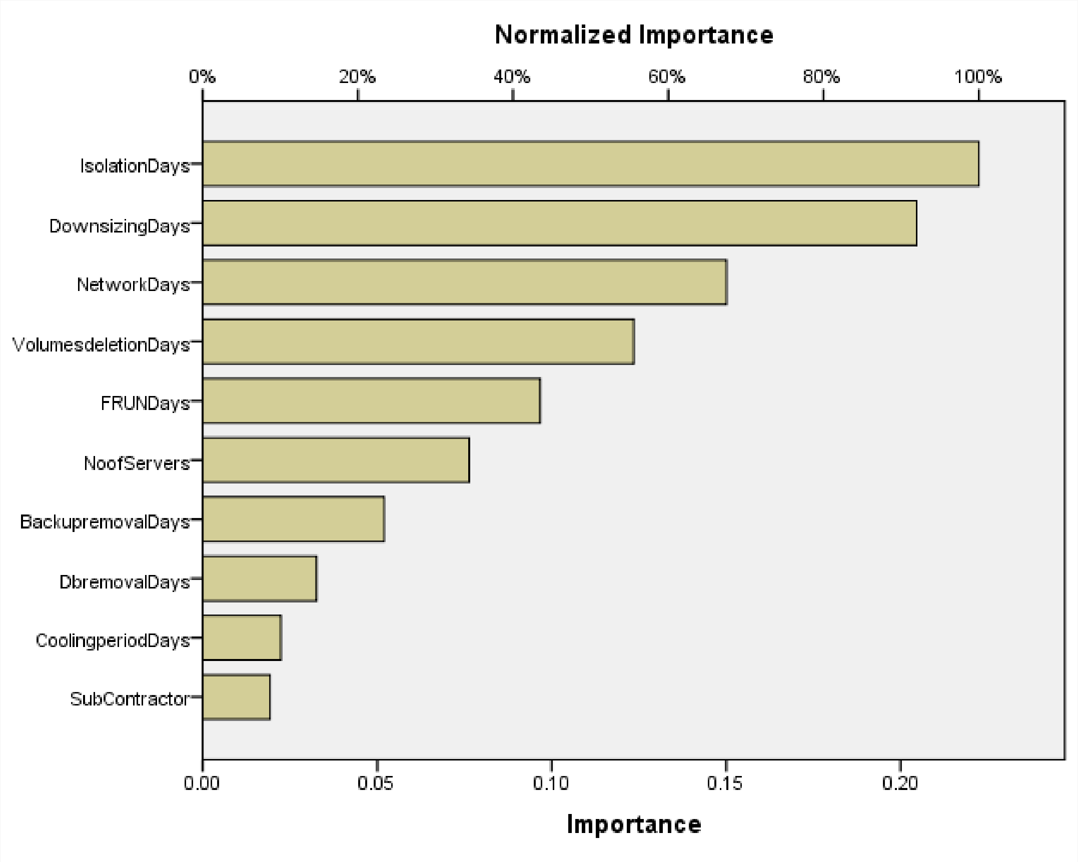
- Decision Tree with Sever Isolation (Days) and Sub-Contractor as key drivers indicating delays.

Decision Tree Output 2



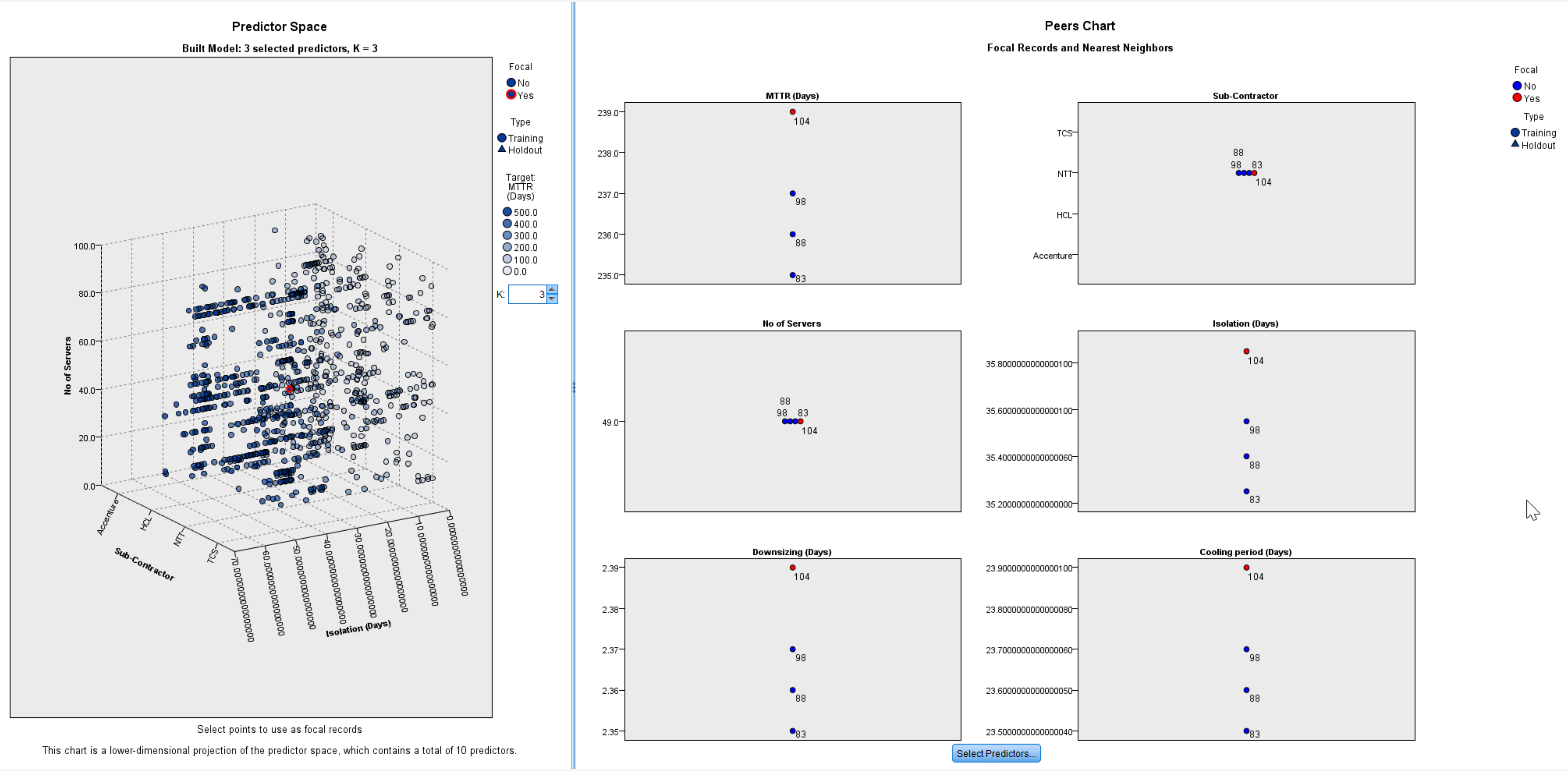
- Decision Tree with Cooling Period (Days) and Sub-Contractor as key drivers indicating delays.
- In case the cooling period is extended beyond 10 days, then the predicted MTTR is more than 100 days.

KNN Classifier Output 1



Normalized Importance using KNN algorithm

KNN Classifier Output 2



KNN model consisting of servers, sub-contractor, and isolation (days)

Model Evaluation

Decision Tree Classifier	
Accuracy	90.30%
Mean Absolute Error	9.67
Min Absolute Error	1%
Max Absolute Error	23%

Metrics for Decision Tree Prediction

K Neighbors Classifier	
Accuracy	96.70%
Mean Absolute Error	3.27
Min Absolute Error	1%
Max Absolute Error	27%

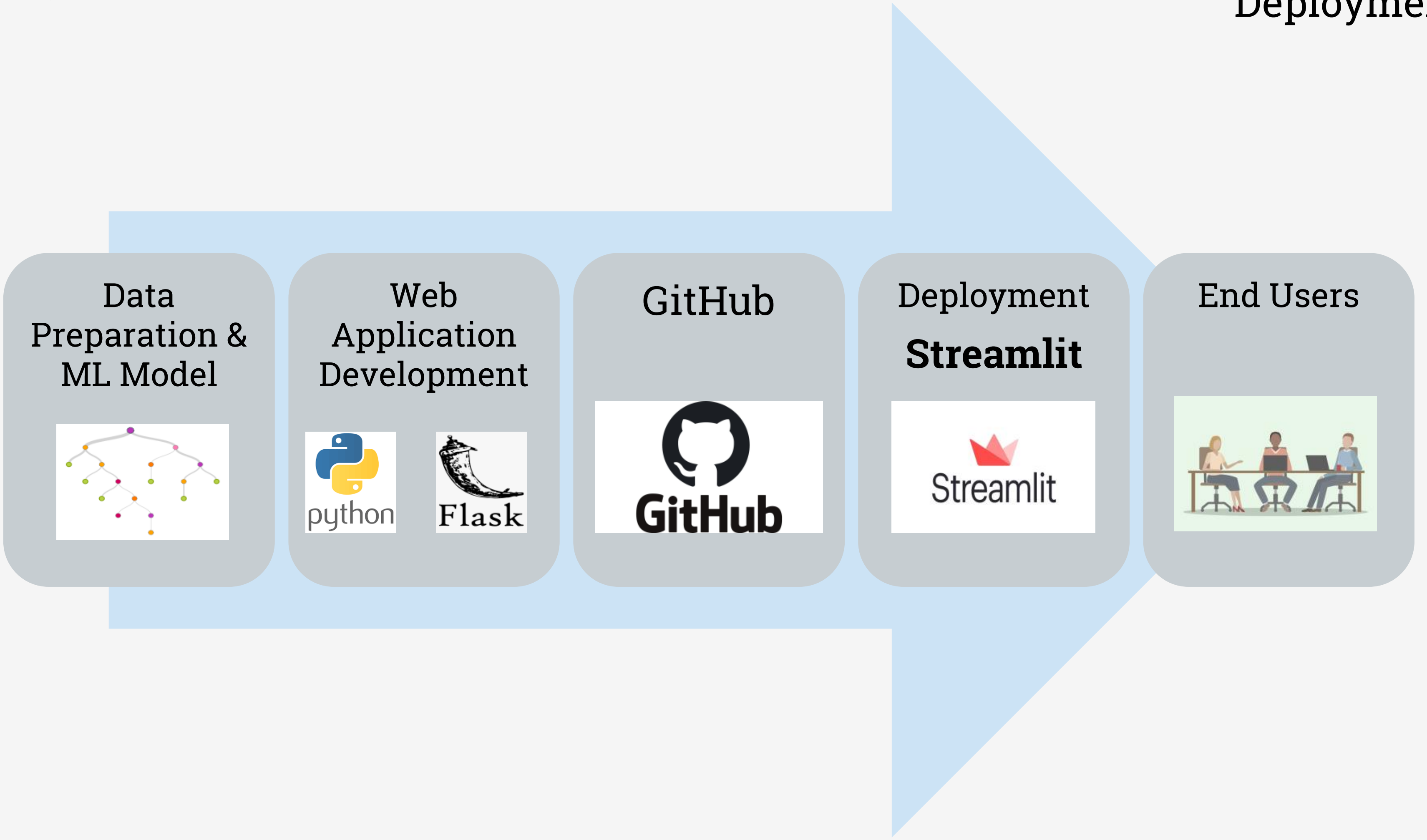
Metrics for K Nearest Neighbors Prediction

- Key drivers influencing the MTTR based on the decision tree are -
  - ✓ Server Isolation (Days)
  - ✓ Cooling Period
  - ✓ Sub-Contractor (Vendor)

- KNN Classifier predicts the MTTR with an accuracy of 96.70%. The model can be used to create an alert mechanism for categorizing new tickets with the predicted MTTR values.

Model Deployment

Deployment Plan



Results and Insights

- Output from two decision trees showcased the following Key Drivers influencing the MTTR –
  - ✓ Isolation (Days)
  - ✓ Cooling Period
  - ✓ Sub-Contractor (Vendor)

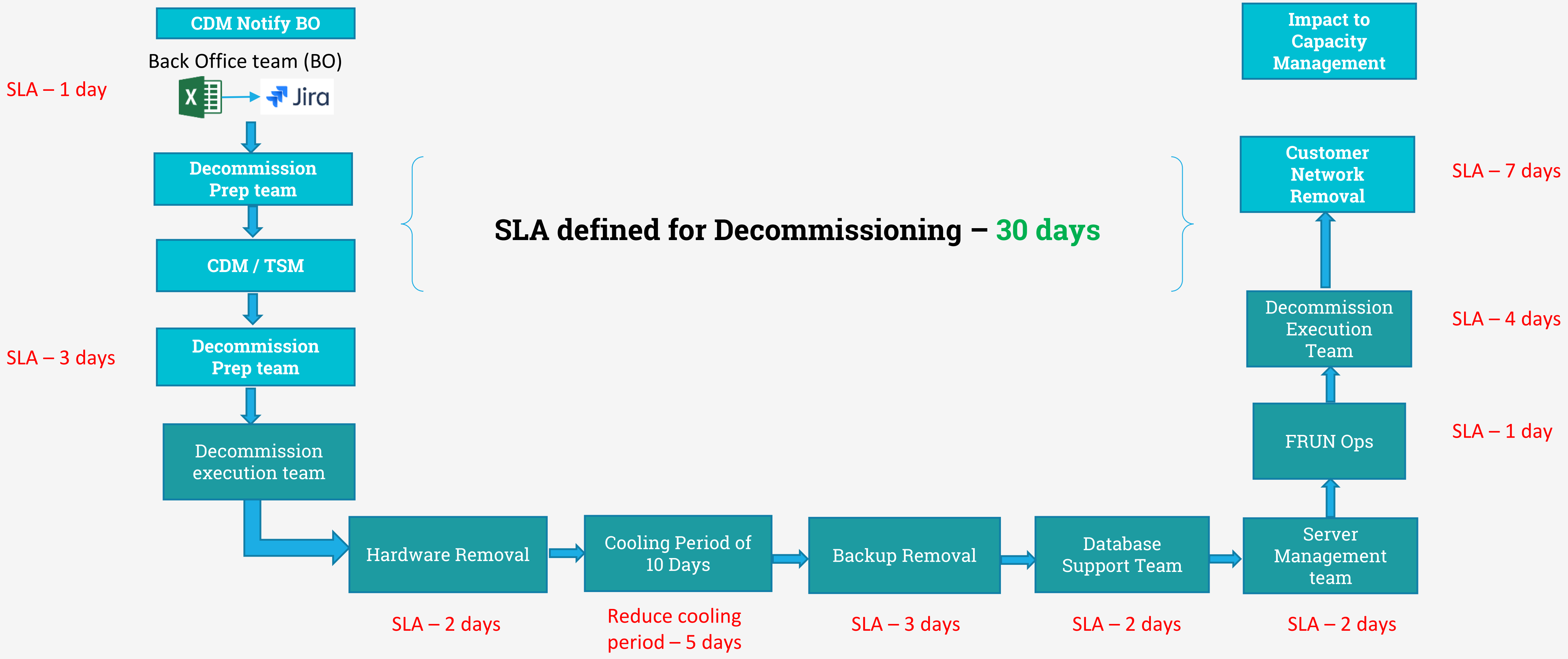
- If the server isolation is completed between 2.7 – 6.7 days, then the predicted MTTR is 66 days. If the server isolation is performed in less than 2.7 days, then the predicted MTTR is 9 days.

- Another decision tree considered the Cooling Period as a key input parameter. If the cooling period is between 1 – 8 days, then the predicted MTTR is 31 days. If the cooling period is >8 days, then the predicted MTTR is 129 days.

- MTTR for Sub-Contractor showed huge delays as the server being isolated on an average is between 12.9 to 19.3 days indicating an MTTR of 101 days being achieved.

REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

End to End Decommission (TO-BE) Process



Conclusion

- The study aims to establish a link between delays and key drivers identified in the process.
- Decision Tree is able to predict the factors that contributes towards the delay.
- Costs will be reduced along with delays with the SLAs being implemented in the process.

Future Work

This project does not cover vendor analysis. In the future study, the scope of vendor analysis is to identify if the type of vendor can be a key driver causing the delays in decommissioning.

Variables such as “Staff Competency”, “Absenteeism rate”, “Number of staff employed”, and “staff experience” shall be captured in future studies that can help provide recommendations to vendors for optimizing the decommissioning process.

Real-time dashboard will be created for server decommissioning as only system and customer decommissioning dashboards are available.

REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956, Approved by AICTE, New Delhi

References

1. Charbuty, B., & Abdulazeez, A. (2021). Classification Based on Decision Tree Algorithm for Machine Learning. *Journal of Applied Science and Technology Trends*, *2*(01), 20–28. https://doi.org/10.38094/jastt20165

2. Radioactive Waste Management, N. (2018). *Preparing for Decommissioning During Operation and After Final Shutdown Radioactive Waste Management and Decommissioning 2018*.

3. Jiao, S. R., Song, J., & Liu, B. (2020). A Review of Decision Tree Classification Algorithms for Continuous Variables. *Journal of Physics: Conference Series*, *1651*(1). https://doi.org/10.1088/1742-6596/1651/1/012083

4. Kang, H. (2013). The prevention and handling of the missing data. *Korean Journal of Anesthesiology*, *64*(5), 402–406. https://doi.org/10.4097/kjae.2013.64.5.402
Krishna M V. (2016). *asset-data-linkage-using-artificial-intelligence-method-of-machine-learning-2016*.

5. Nyarko-Boateng, O., Adekoya, A. F., & Weyori, B. A. (2020). Using machine learning techniques to predict the cost of repairing hard failures in underground fiber optics networks. *Journal of Big Data*, *7*(1). https://doi.org/10.1186/s40537-020-00343-4
Quiry, P., Vernimmen, P., Fur, Y., Dallocchio, M., & Salvi, A. (2009). *Corporate Finance: Theory and Practice*.

6. SIMS Lifecycle Services. (2020a). *SOLUTIONS FOR GLOBAL DATA CENTER DECOMMISSIONING*.

7. SIMS Lifecycle Services. (2020b). *SUSTAINABLE DATA CENTER DECOMMISSIONING*.

8. Song, Y. Y., & Lu, Y. (2015). Decision tree methods: applications for classification and prediction. *Shanghai Archives of Psychiatry*, *27*(2), 130–135. https://doi.org/10.11919/j.issn.1002-0829.215044

9. Vuttipittayamongkol, P., Tung, A., & Elyan, E. (2021). A Data-Driven Decision Support Tool for Offshore Oil and Gas Decommissioning. *IEEE Access*, *9*, 137063–137082. https://doi.org/10.1109/ACCESS.2021.3117891

10. W, B. A., A, K. K., & B, A. D. (2018). Issues with data and analyses: Errors, underlying themes, and potential solutions. *Proceedings of the National Academy of Sciences*, *115*(11), 2563–2570. https://doi.org/10.1073/pnas.1708279115

11. Zeng, S., Melville, P., Lang, C. A., Boier-Martin, I., & Murphy, C. (2008). Using Predictive Analysis to Improve Invoice-to-Cash Collection. *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1043–1050. https://doi.org/10.1145/1401890.1402014

12. Zhang, X., & Jiang, S. (2012). A splitting criteria based on similarity in decision tree learning. *Journal of Software*, *7*(8), 1775–1782. https://doi.org/10.4304/jsw.7.8.1775-1782

Annexure

Plagiarism score

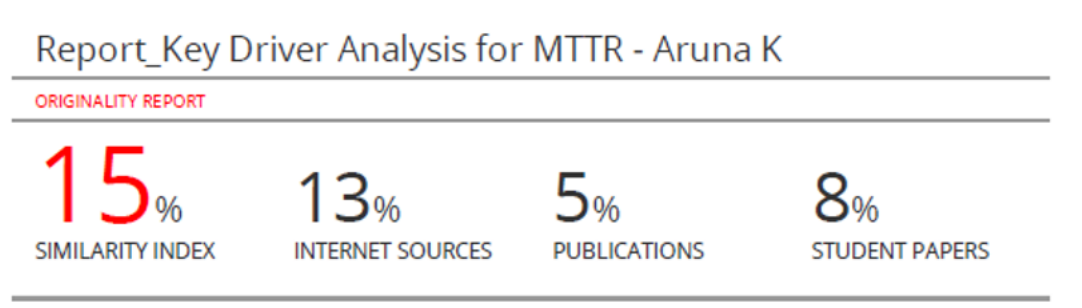
This is to certify that this project report titled **Key Driver Analysis for MTTR using Machine Learning Approach** was scanned for similarity detection. Process and outcome are given below.

Software Used: **Turnitin**

Date of Report Generation: **12.08.2022**

Similarity Index in %: **15%**

Total word count: **10805**



REVA
UNIVERSITY
Bengaluru, India

Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

Thank you!