Corporate Internship
ExcelFore - Corportion

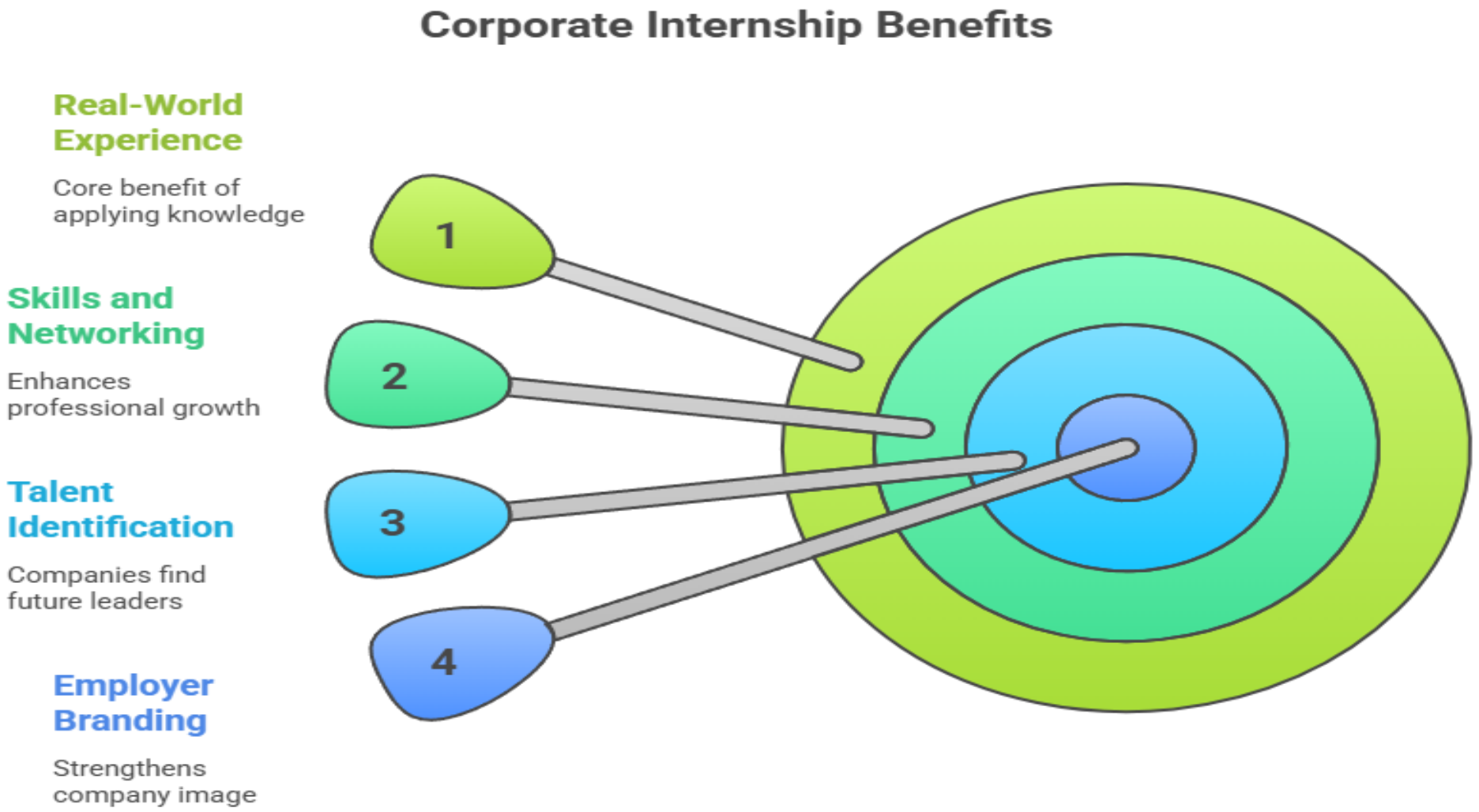
Name of the Student : Nikhil P Marihal
SRN of the Student : PES1UG23BB665

Name & Designation of the College Guide :
Prof Nitish Rajamane

Name & Designation of the Company Guide: Mr. Gurudath
Principle Engineer

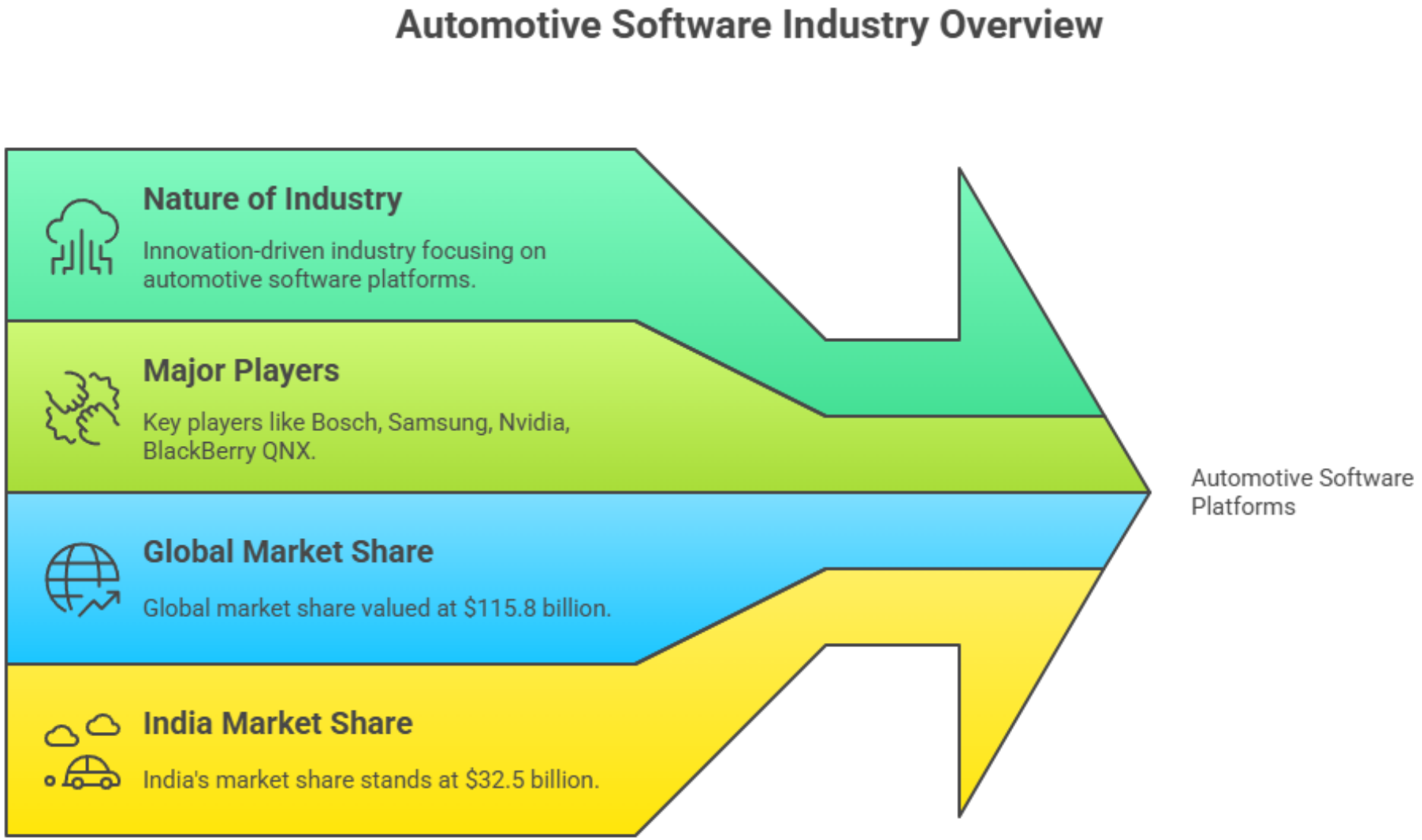
Corporate Internship

Introduction of corporate internship



Corporate Internship

Industry Profile-Size, Nature, Major players, Related Industries, Market Share



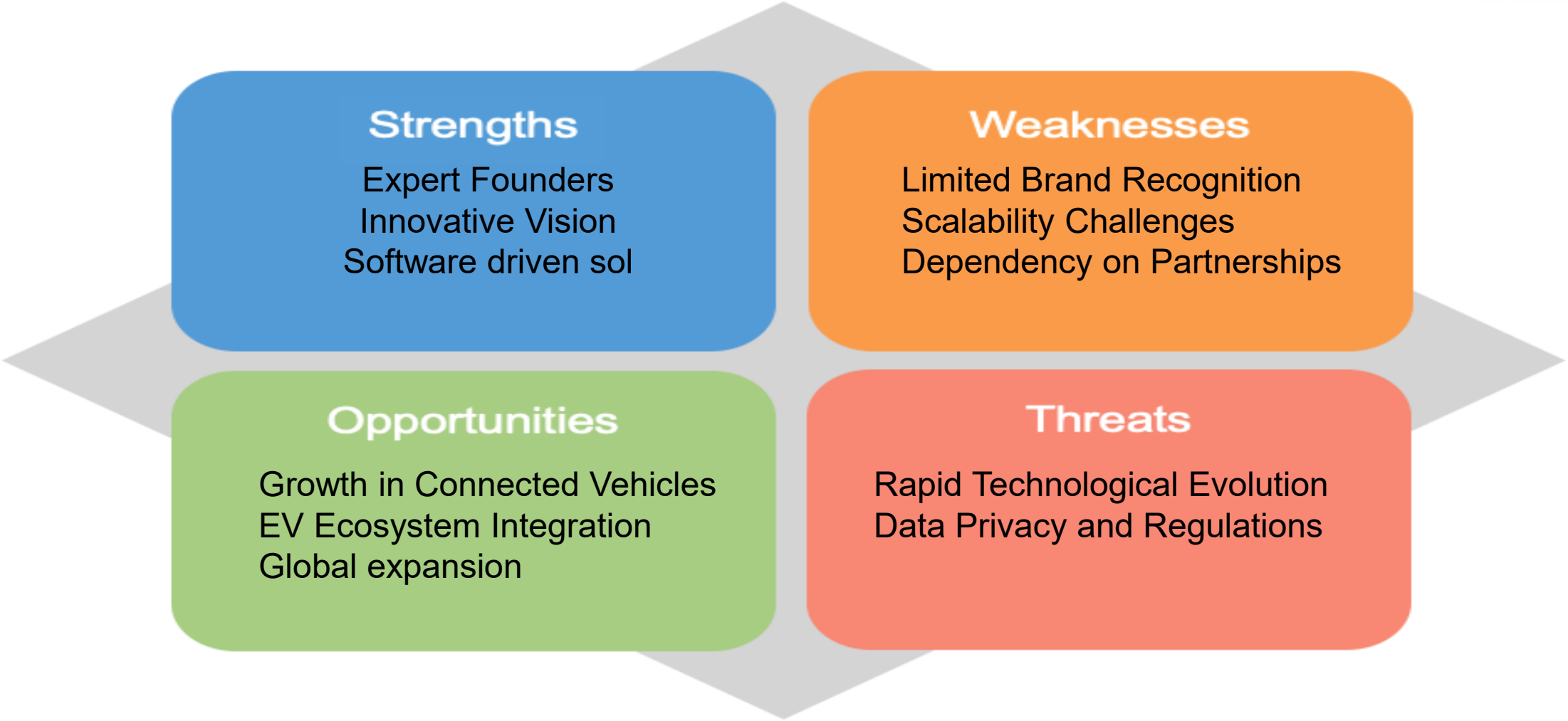
Corporate Internship

Company Profile-Details of the Organization, Organization structure etc.

Vision: Driving Connectivity | Enabling Innovation | Shaping the Future of Automotive Tech.
Background: Founded to transform automotive connectivity by integrating vehicles with the cloud. Pioneered seamless cloud-vehicle integration for software-defined, customer-centric cars.

Corporate Internship

SWOT analysis



Corporate Internship

2 Functional departments worked for 2 weeks

- Machine Learning
- Finance

Organizational Challenges Prioritization

Integration Complexity

Integration complexity poses significant challenges with minimal impact.

Security and Compliance

Security and compliance are critical but complex challenges.

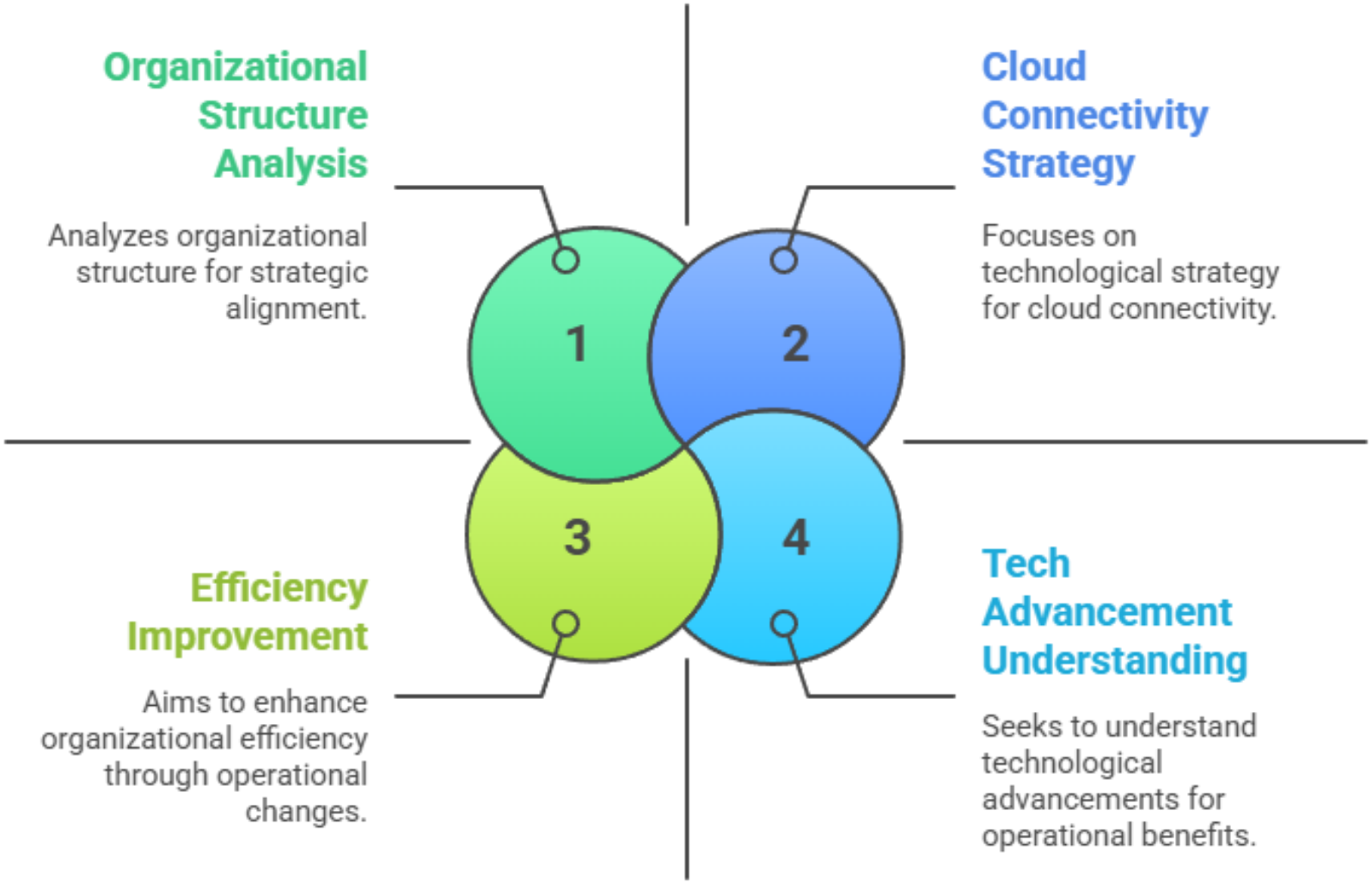
Customization

Customization is straightforward but has limited organizational impact.

Scalability Issues

Scalability issues are impactful yet relatively simple to address.

Strategic Study Objectives and Needs



Executive Summary: This project focused on developing a machine learning model to predict loan defaults by analyzing a dataset of banking clients. The initial exploratory data analysis identified key client attributes and, most critically, revealed a significant class imbalance, with far more non-defaulters than defaulters. This imbalance poses a major challenge as it can lead to biased models that are poor at identifying high-risk clients

Corporate Internship

Article Writing

- **Background:** While lending is a primary revenue source for banks, loan defaults pose a significant financial risk. This project analyses client data to identify key factors that predict the likelihood of a default. The objective is to build a more robust risk assessment model, enabling smarter lending decisions and minimizing potential losses for the institution.

- **Problem Evaluation:** The primary challenge is building a predictive model for loan defaulters using a dataset with a significant class imbalance, where non-defaulters greatly outnumber defaulters. This issue can bias the model, making it ineffective at identifying the crucial minority class (defaulters) and thus unreliable for risk management. Mitigating this imbalance is a critical step.

- **Suggestions and Recommendations:**

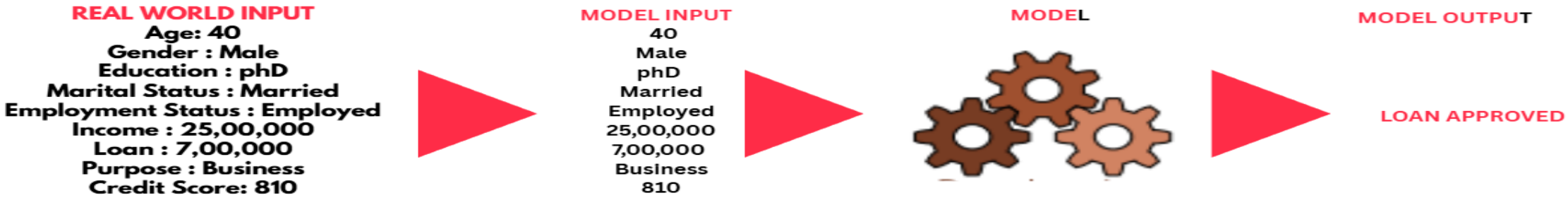
1. **Resampling**: Overfitting, Underfitting
2. **Model Level Approaches**: Algorithms, Cross-Validation
3. **Focus on Key Metrics**: Prioritize Precision, Recall and F-1 Score

- **Conclusion:** Accurately predicting loan defaults is a critical task for risk management in the banking sector, but imbalanced datasets often challenge it. By creating a more equitable dataset for training, we can have the necessary groundwork for developing an unbiased and effective machine learning model.

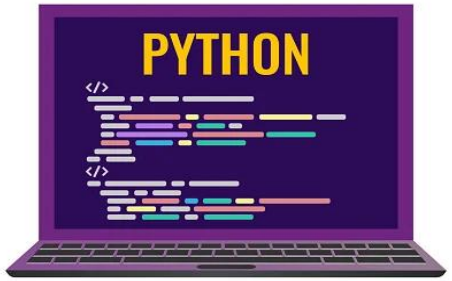
Corporate Internship

Analysis

- **EDA**: Initial Inspection, Descriptive Statistics, Duplicate Check
- **Statistics**: Cross-tabulation, Aggregation, Binning
- **Visualization**: Categorical Analysis, Distribution Analysis, Histplot, Value Counts.
- **Visual Presentation of the working model**:



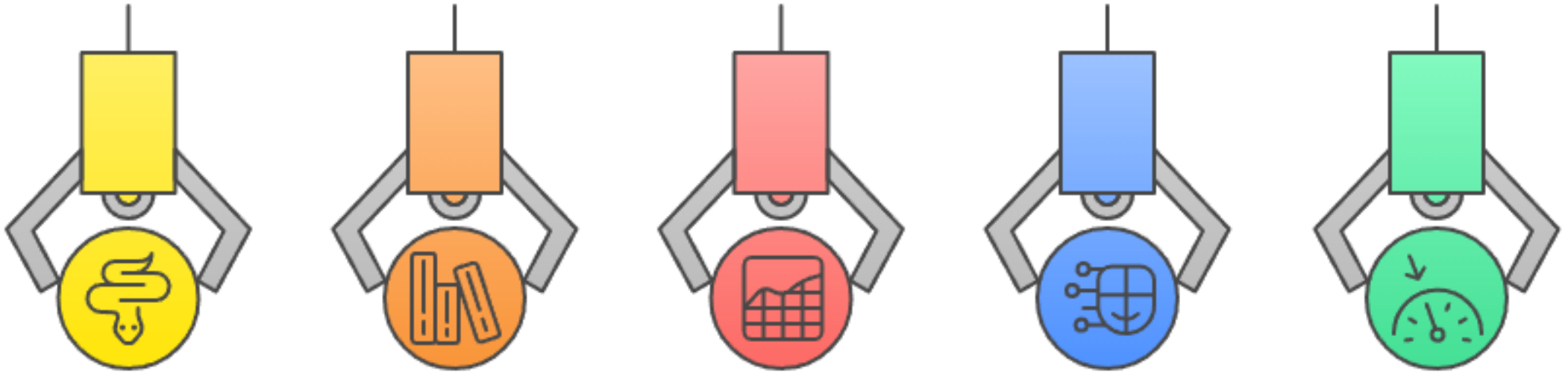
Link to Analysis: Click on the Python Image to view the codes.



Corporate Internship

Learning Experience

Key areas of learning



Python Essentials

Fundamental data structures and object-oriented programming in Python.

Key Libraries

Essential Python libraries for data manipulation and analysis.

Statistics & Analysis

Statistical concepts and techniques for data analysis.

Machine Learning

Core machine learning paradigms and algorithms.

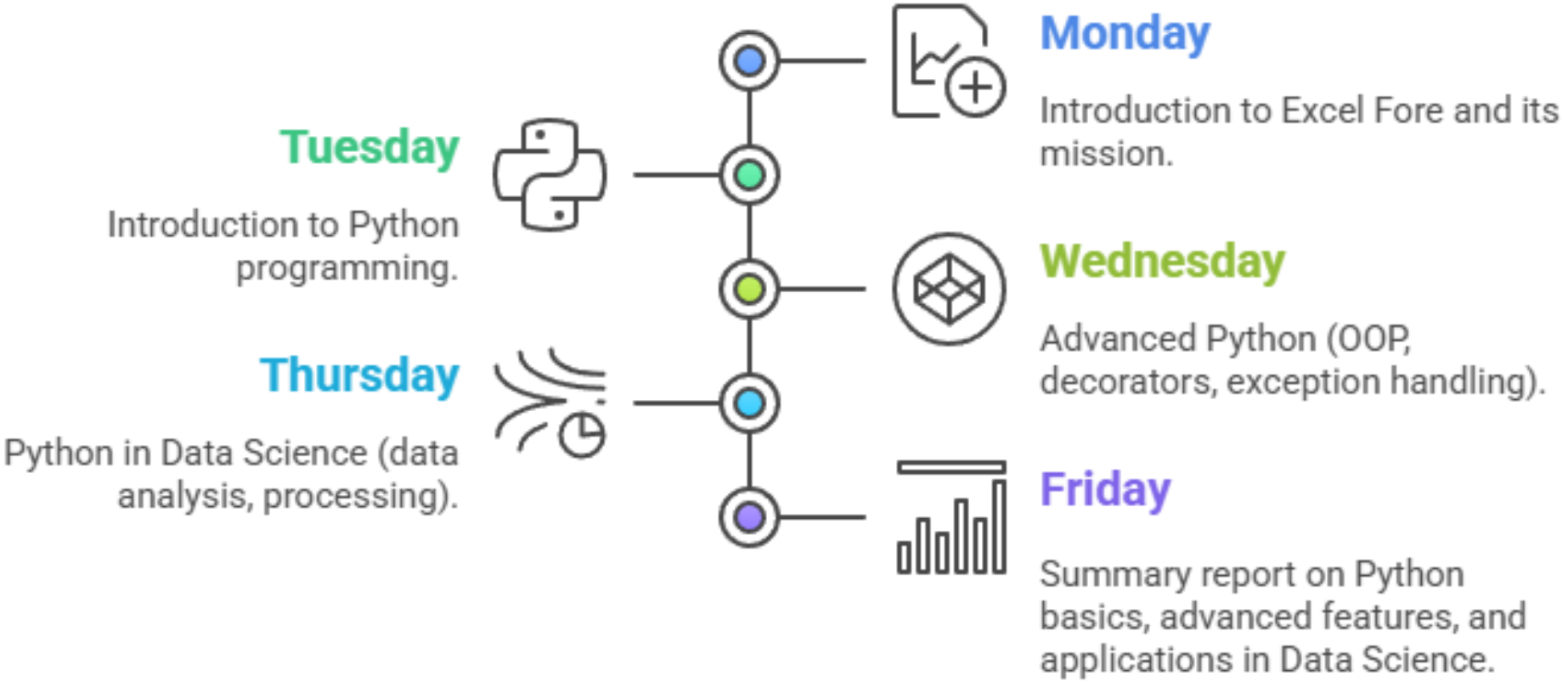
Modeling & Tuning

Techniques for building, refining, and validating machine learning models.

Corporate Internship

Week 1 Report

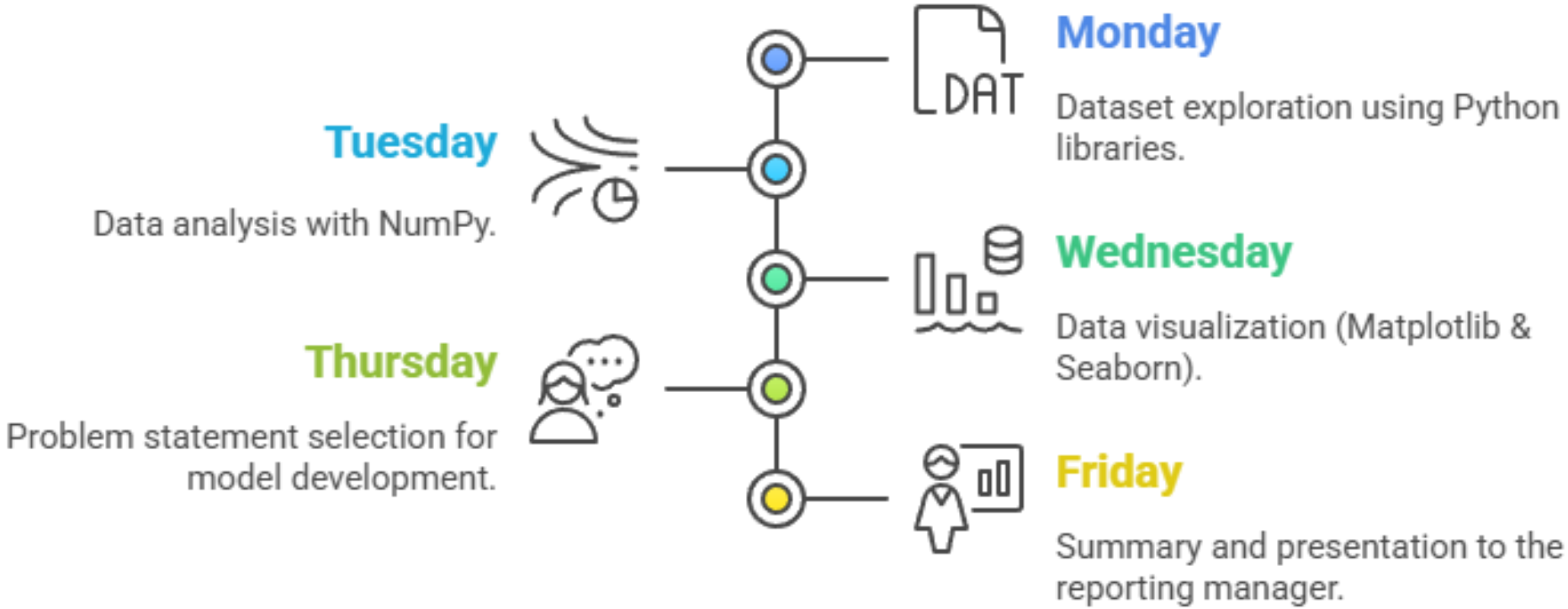
Week-long Journey Through Excel and Python



Corporate Internship

Week 2 Report

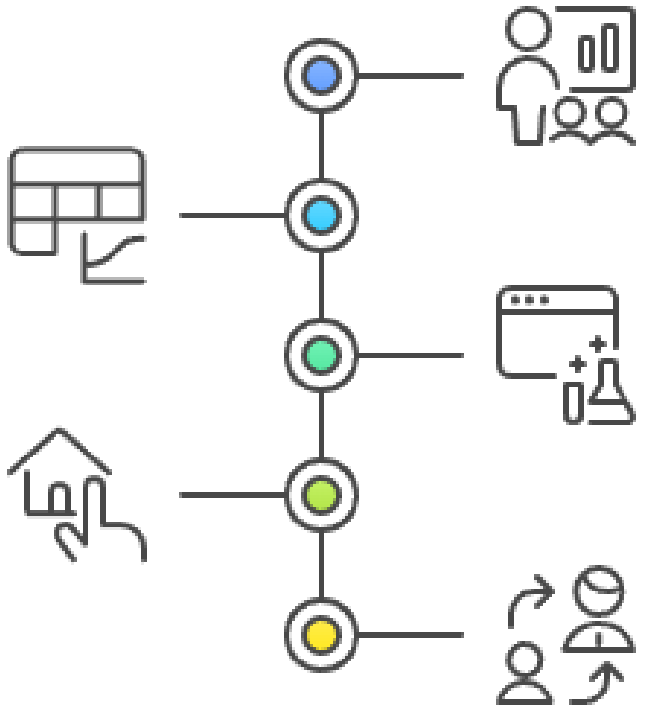
Weekly Data Analysis and Presentation Timeline



Corporate Internship

Week 3 Report

Weekly Statistics Learning Journey



Monday

Introduced to Statistics & Statistics Visualization

Tuesday

Analysing Data & Data Distribution

Wednesday

Probability & Hypothesis Testing

Thursday

Feature Selection & Engineering , Model Selection

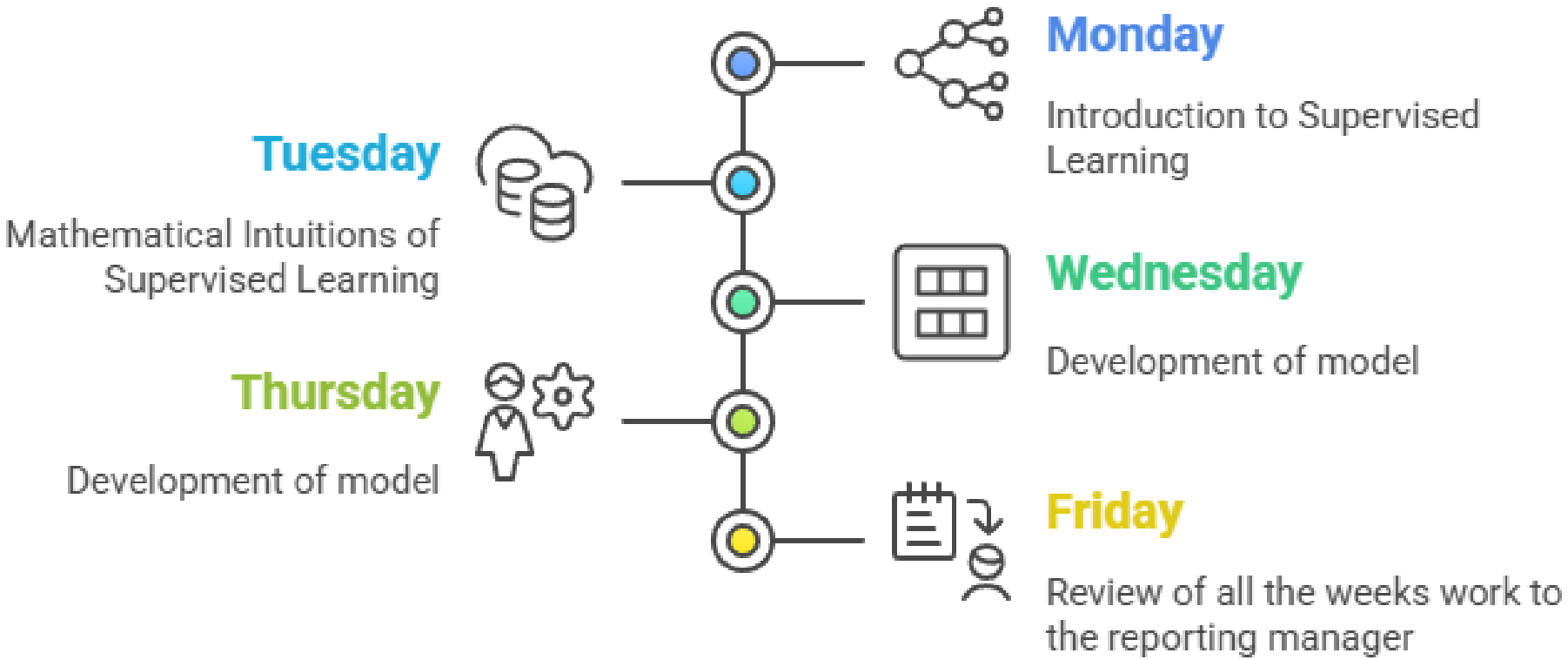
Friday

Review of all the weeks work to the reporting manager

Corporate Internship

Week 4 Report

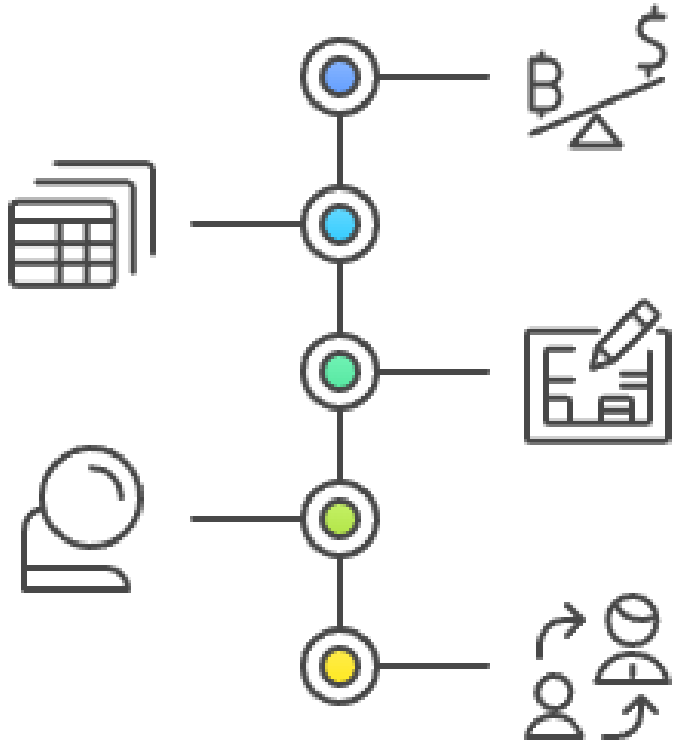
Weekly Learning Journey in Supervised Learning



Corporate Internship

Week 5 Report

Weekly Data Imbalance Management Timeline



Monday

Dealing with imbalance data

Tuesday

Types of imbalance data

Wednesday

Development of model

Thursday

Development of model

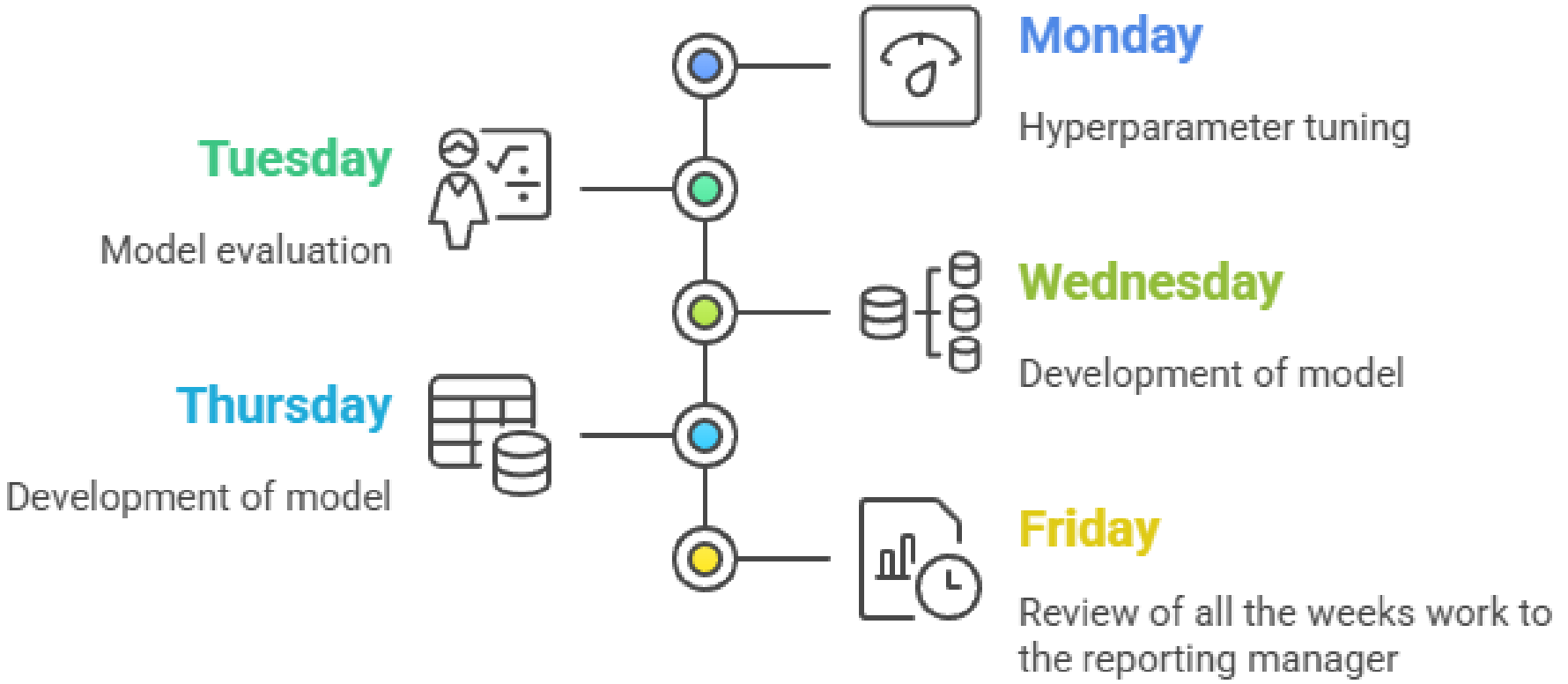
Friday

Review of all the weeks work to the reporting manager

Corporate Internship

Week 6 Report

Weekly Machine Learning Project Timeline



THANK YOU

Name of the Student : Nikhil P Marihal
SRN of the Student : PES1UG23BB665