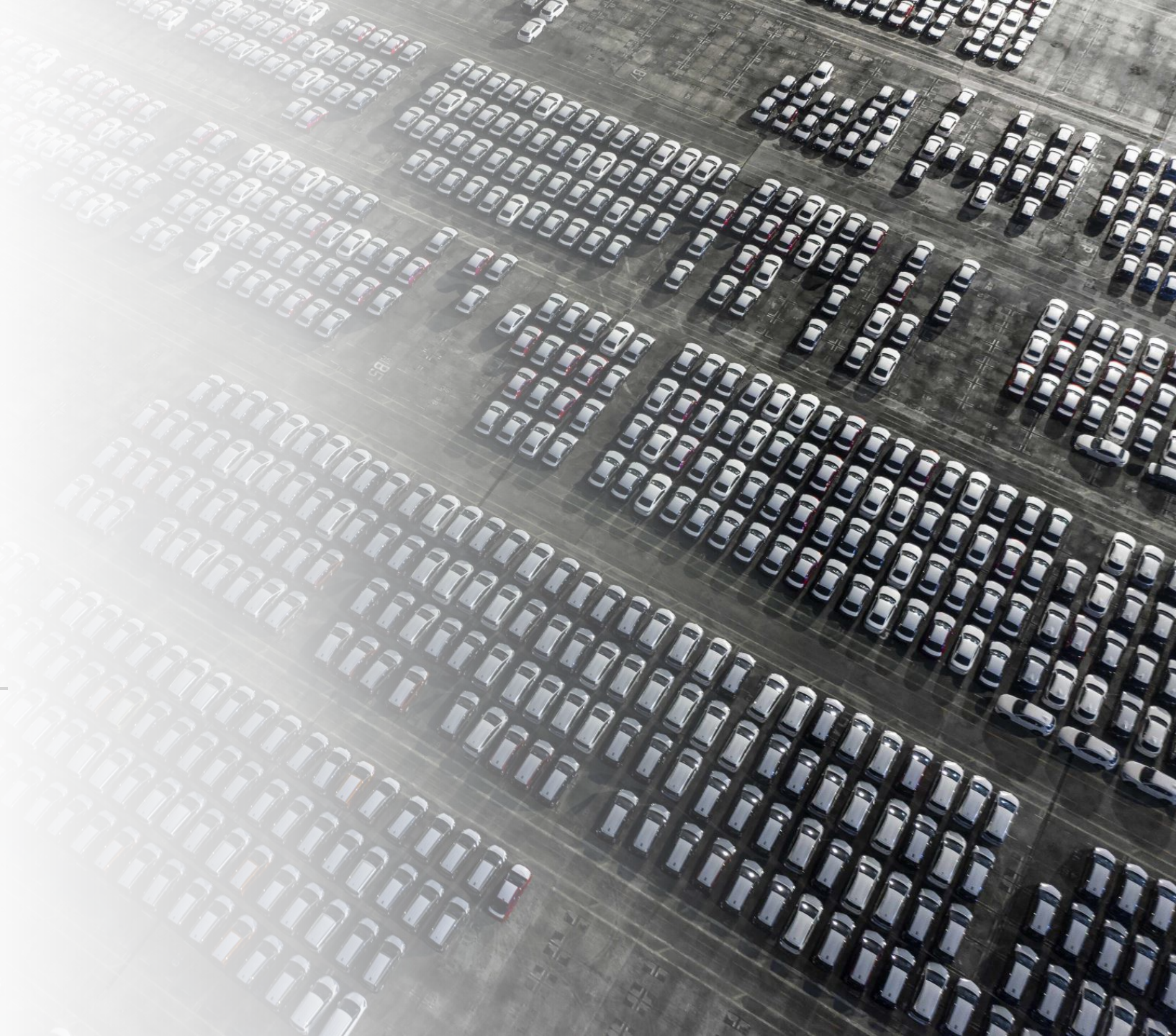
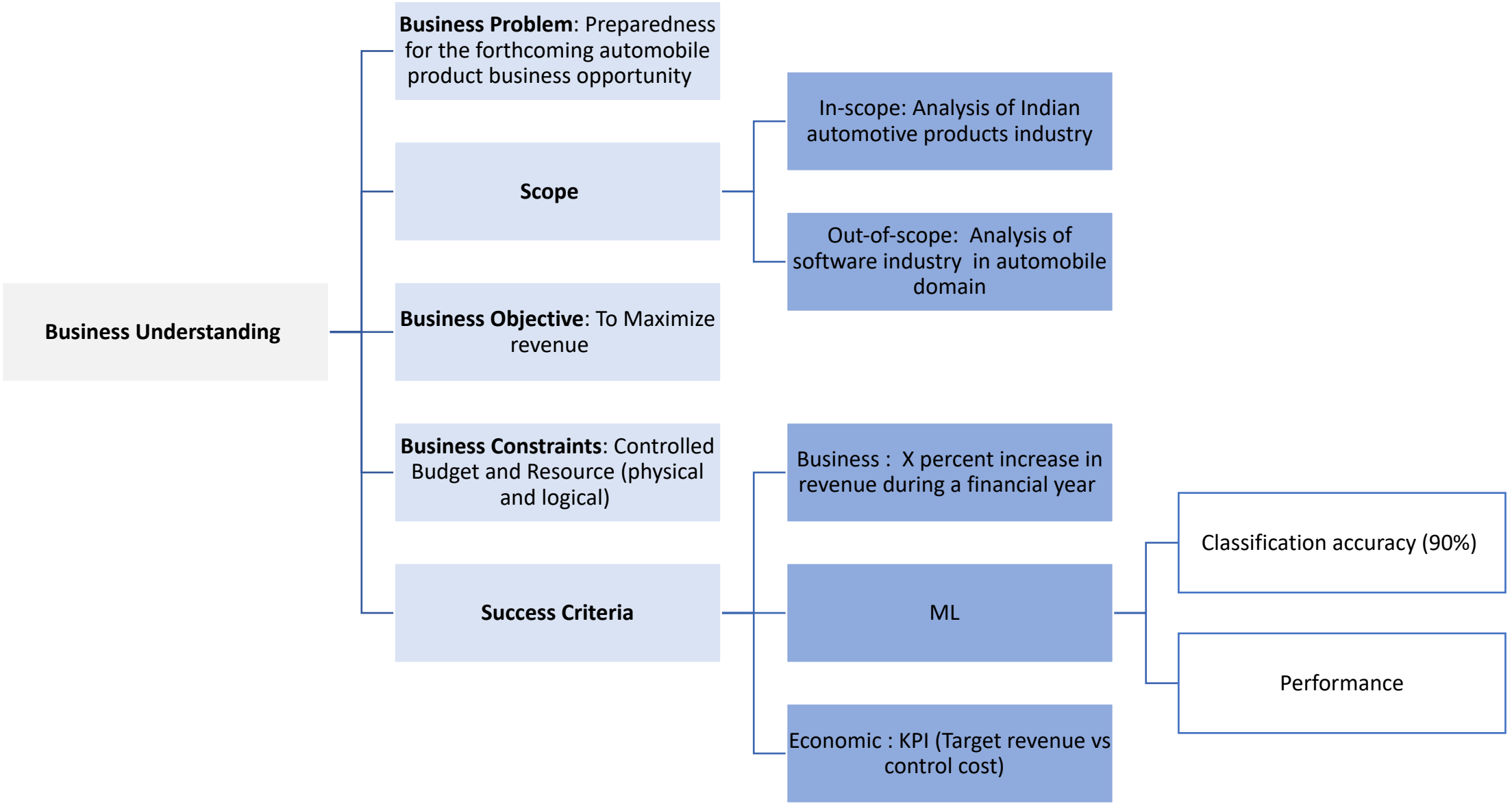
ML: Automobile Product Business Opportunity Prediction



Group1

- Ajith Sanikommu - 12120006
- Anjana Rajan - 12120080
- Bhargavi Peddapati - 12120067
- Rohini Singh - 12120059
- Shantanu Srivastava - 12120061

Business and Data Understanding



ML Feasibility

- Predictive features: Positive, negative and neutral sentiments, Four-wheeler sales, on automotive parts sales turnover
- Data source: Amazon to capture sentiments based on review and rating and ACMA (Automotive Component manufactures of India) for overall automotive industry turnover
- Deployment: On local inference system
- Legal constraints: NA
- Feasibility study outcome
    - Correlation between consumer’s sentiments and automotive industry sales for the give period (FY2019-FY2021)

Input & Output Parameters

- **Input Parameters:**
  - Four-wheeler sales
  - Positive Reviews
  - Negative Reviews
  - Neutral Reviews

- **Output Parameters :**
  - Sales Growth

Data Understanding

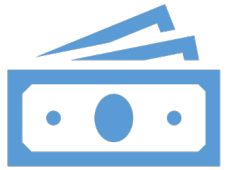
- **Version Control** : NA
- **Quality Verification**:
  - Data description
    - Sentiments of review comments extracted from Amazon
    - Sales turnover from ACMA
  - Data requirements
    - data size is approximately 2300 records
  - Data verification :
    - Authenticity of source
    - Data sanity

Success Criteria

Business : X percent increase in revenue during a financial year

ML

- Classification accuracy (90%)
- Performance
  - Robust
  - Scalable
  - Complexity
  - Resource required
- Explainability: Post hoc explainability

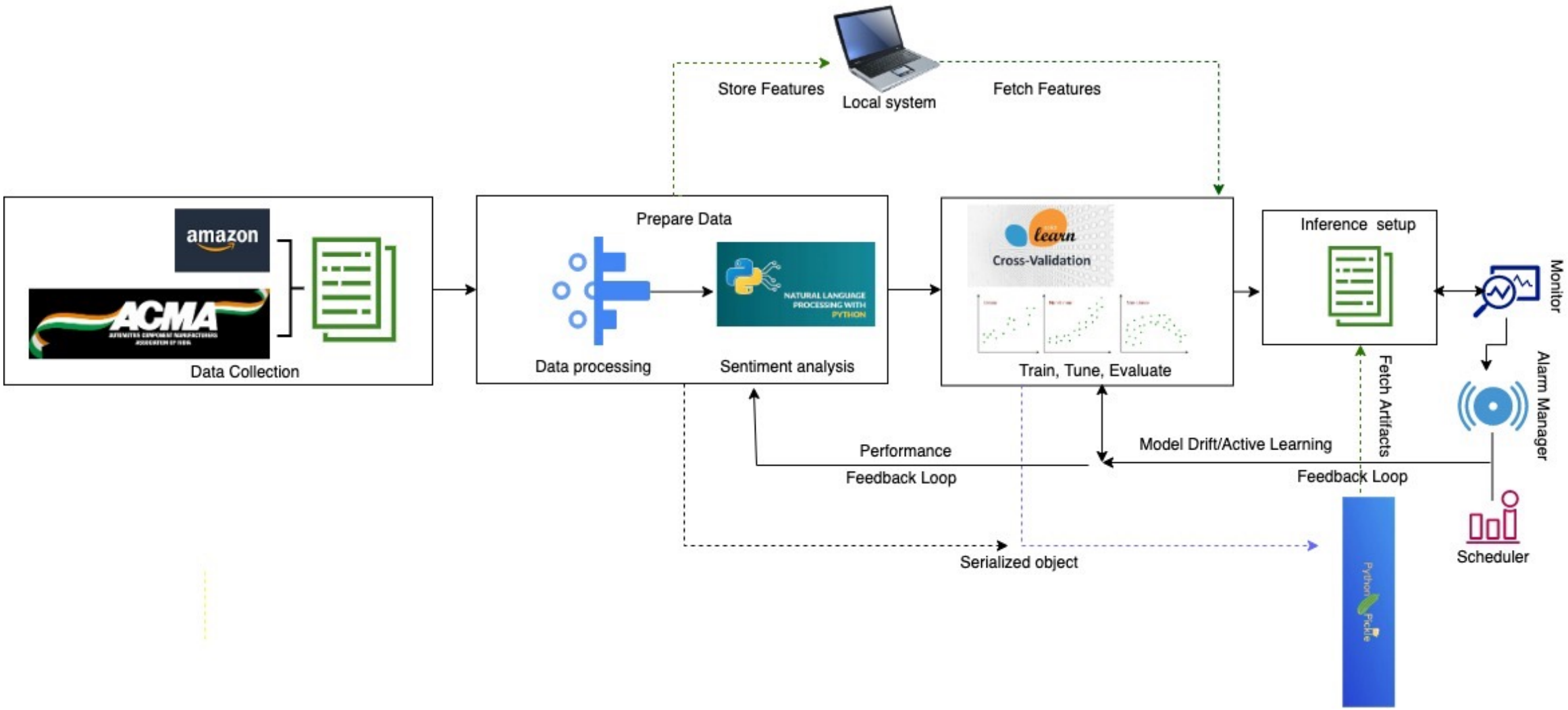


Economic : KPI (Target revenue vs control cost)

Modeling

- **Model Selection and building**
  - NLP classification (using product review data to generate sentiment insights)
  - Correlation analysis among deduced sentiments, four-wheeler sales, and automobile components sales turnover
  - Regression between sales and identified feature (positive sentiments and four-wheeler sales)
  - Train on FY2019 to FY2021 data, and on inference setup we have evaluated model with FY2022 data for the prediction
  - Model evaluation based on R^2

ML- Architecture



Evaluation

Explainability	Factors such as positive , negative and neutral sentiments which have direct influence on ML prediction
Result Assessment	Model validated on test data - Compliance accuracy >= 98% - Revenue visualization as per business objective - Economic success: Higher Revenue to Cost ratio

Deployment & Post-Production

Deployment Strategy

- Code Maintains at GitHub Repository
- Deploy ML model using Flask
- CI/CD/CT using Buildkite
- Seamless adaptability - Deployment is agnostic to Cloud / On prermise

Monitoring/Maintenance

Monitoring tools such as Neptune, Grafana + Prometheus(Open source) help us to log and store ML metadata and monitor:

- Model training
- Evaluation testing
- Hardware metrics display
- Alerts

Thank You