High Level Plan – Improve the Safety of Breda

High Level Plan - Improve the Road Safety in Breda

Ron L. Tabuchov
Mohamed K. A. M. Elshami
Daria Vladutu
Peter Husen

Data Science and Artificial Intelligence, Breda University of Applied Science

DISCOVER YOUR WORLD

Breda University
OF APPLIED SCIENCES

Index

1   Background

Road safety is a universal concern that impacts public health, urban mobility, and economic efficiency. Breda faces significant challenges with road safety as traffic accidents result in personal injury and disruptions to the urban transport system. By addressing this issue, we can leverage cutting-edge technology to create safer cities.

2   Goal

Make Breda safer by reducing the number of accidents while combining AI and ML to notify the driver about different risk zones in the city.

3   Method

Following the CRISP-DM life cycle along with identifying key risk factors, developing machine learning models, and implementing effective interventions to reduce accidents.

Current methods for identifying danger zones rely on historical data and manual analysis, which can be time-consuming and less effective.
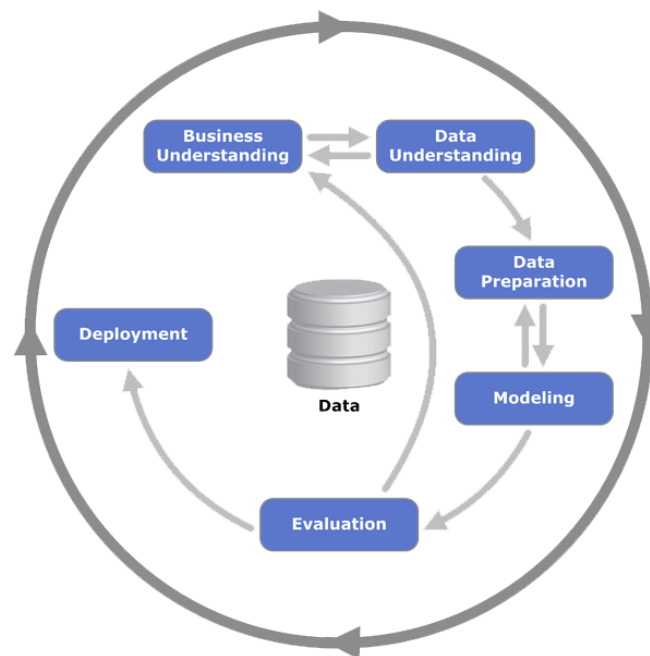

Figure 1. CRISP-DM Life Cycle

4 Tasks

Tasks are divided according to the agreed weekly plan and leads the team to achieve the project's goal and objectives.
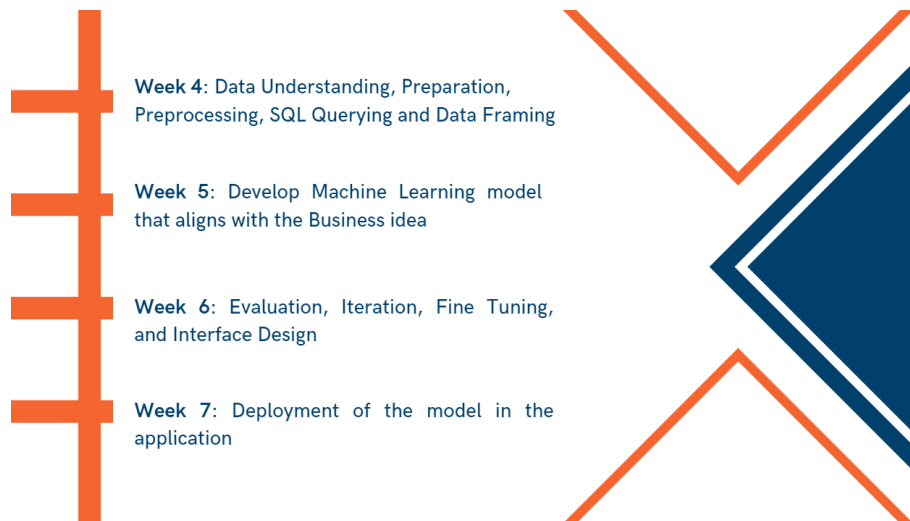
Week 4: Data Understanding, Preparation, Preprocessing, SQL Querying and Data Framing

Week 5: Develop Machine Learning model that aligns with the Business idea

Week 6: Evaluation, Iteration, Fine Tuning, and Interface Design

Week 7: Deployment of the model in the application

Figure 2. Weekly Plan from the Proposal Presentation

Week 4: Data

Monday: Project Planning and Data Understanding

- o Stand-up meeting: Project overview and weekly planning.
- o Load the data to Jupiter notebook.
- o Understand the data sets.

Tuesday: Data Pre-processing

- o Stand-up meeting: Daily planning.
- o Query the data.
- o Define preprocessing tasks and document in PDF file.
- o Preprocess the data set (removes unnecessary columns, handles missing values, addresses any inconsistencies, errors, or outliers).
- o Create tasks in Trello.

Wednesday: Data Pre-processing

- o Stand-up meeting: Daily planning.
- o Perform standardizes formats and encodes categorical variables.
- o Join, normalize, and standardize numerical features.
- o Splits the dataset into training, validation.
- o Extracts relevant features or creates new features from date and time information.
  OR
  creates new features from existing features to generate more informative ones.

Thursday: Data Confirmation

- o Stand-up meeting: Daily planning.
- o Team re-view for data preprocessing.
- o Perform adjustment (if necessary).
- o Data confirmation with mentor (green light approval for ML modelling).
- o Evidencing in Jupiter notebook, Work & Learning Log.

Friday: Evaluate

- o Stand-up meeting: Future planning.

Week 5: Machine Learning & Legal Framework

Monday: Modelling & Risk Assessment

- o Stand-up meeting: Weekly & Daily planning.
- o Make random forest & another model.
- o Legal framework – identify the level of risk and explains it.
- o Train the models.
- o Evaluate and compare outputs (Markdown for first results).
- o Identify framework vital legal requirements and obligations.

Tuesday: Hyperparameter & Framework

- o Stand-up: Discuss random forest model with team.
- o Model iteration.
- o Establish legal framework.
- o Provide feedback: Legal framework & Selection of ML models.
- o Create iteration 2 (use different hyper-parameters for both models).
- o Review Legal framework after iteration 2.
- o Select suitable frameworks.
- o Add reasoning for not including legal frameworks. argue why things are included or excluded based on the facts and details (examples and sources).

Wednesday: XAI & Legal

- o Stand-up meeting: Discuss model performance.
- o Analysing trade-off (XAI).
- o Define performance metrics (e.g., accuracy, precision, recall, F1 score) and criteria for model complexity (e.g., training time, interpretability).
- o Markdown conclusions from the model comparison.

Thursday: Catching up.

- o Stand-up meeting: Discuss model performance.
- o Feedback / Review Legal framework.
- o Feedback / Review ML.

Friday: Evidencing

- o Stand-up meeting: Future planning.
- o Catching up / complete delayed tasks.
- o Evidencing & Documenting ILO 6.0 & 3.0

Week 6: Deployment & Interface Design

Monday: Interface Design & Deployment

- o Stand-up: Daily planning.
- o Review and catch up with tasks of week 5.
- o Interface Design – Create Figma 2 version.
- o Provide first feedback.
- o Write A/B test.
- o Publish A/B test.

Tuesday: A/B test & Enviroment

- o Stand-up: Daily planning.
- o Collect responses (minimum 15 per version).
- o Analysis A/B test.
- o Update and adjusts A/B Figma version.
- o Select application demo.
- o Create demo application.
- o Deliver Python scripts & review scripts with peer.
- o Create a virtual environment.
- o Dependency management (Create a requirements.txt file listing all dependencies).
- o Utilize the virtual environment alongside a package manager.

Wednesday: Code cleaning

- o Code cleaning according to industry standards (logging, linting, doc strings, code formatting, refactoring).
- o Set up logging framework ('pylint', 'flake8')
- o Write docstrings for all functions and classes (code formatter 'black' ensures consistent code style).
- o Final code review.

Thursday: Unit-test

- o Write unit-test for the code using 'pytest'.
- o Measure test coverage using 'coverage.py' (30% coverage is required).
- o Install Sphinx for generating documentation.
- o Use docstrings to document code with Sphinx.
- o Project wrap-up meeting & Final template review.

Friday: Interface Design

- o Stand-up: daily planning.
- o Self-evidencing (Work & Learning Log).
- o Presentation preparation
- o Re-view final presentation.

Week 7: Deployment & Presentation

Monday: Individual Contribution

- o Stand-up meeting: Week and daily planning.
- o Confirm completed tasks (Assessment rubric)
- o Completing delayed tasks.
- o Evidence Individual Contribution

Tuesday: Industrial Code

- o Stand-up meeting: Daily planning.
- o Preparing Final Presentation.
- o Completing delayed tasks.

Wednesday: Unit Test & Final Review

- o Stand-up meeting: Daily planning
- o Completing delayed tasks.

Thursday: Evidencing & Presentation

- o Final stand-up meeting
- o Completing delayed tasks.

Friday: Final Project Day

- o Final project presentation and deliverables submissions.

5   Risks

Server Connection Issues:

Risk: The project relies on accessing a server at http://194.171.191.226:3078/y1d/<STUDENTNUMBER>/lab. Any downtime or connection issues with this server could hurt the progress.

Mitigation: Ensure regular backups of data and progress. Have a local development environment set up to continue work offline if needed.

Work and University Time Management:

Risk: Balancing project tasks with personal commitments might lead to delays or reduced quality of work.

Mitigation: Utilize time management tools and techniques. Set clear, achievable goals for each week and stick to a consistent schedule. Use tools like Trello for efficient task management.

Data Quality and Preprocessing Challenges:

Risk: Incomplete or poor-quality data can lead to inaccurate model predictions and results.

Mitigation: Implement thorough data validation and cleaning procedures. Document each step of the preprocessing to ensure transparency and reproducibility.

Coordination and Communication:

Risk: Miscommunication among team members can lead to duplicated efforts or overlooked tasks.

Mitigation: Schedule regular stand-up meetings and ensure all team members are on the same page. Use SharePoint and Teams for clear communication and document sharing.

Technical Challenges with Machine Learning Models:

Risk: Selecting and tuning the wrong models can result in ineffective solutions.

Mitigation: Perform extensive model evaluation and comparison. Document the reasoning for model choices and hyperparameters.

Compliance with Legal Frameworks:

Risk: Failing to adhere to legal requirements can cause ethical and legal issues.

Mitigation: Integrate legal framework checks into the model development process. Regularly review compliance requirements and document adherence.

6   User Story

User: Daria Vladutu, Data Analyst

Goal: To ensure that all modifications and preprocessing steps in the datasets are well-documented.

Benefit: Ensures that all data changes are traceable, reproducible, and transparent, facilitating easier reviews and audits.

User: Ron Lev Tabuchov, Machine Learning Specialist

Goal: To establish structured protocols for model tuning and performance evaluation.

Benefit: Enhances the accuracy and efficiency of predictive models, ensuring that the project delivers high-quality outcomes.

User: Peter Husen, Compliance Officer

Goal: To monitor and integrate compliance with legal frameworks throughout the development process.

Benefit: Keeps the project aligned with legal standards and ethical considerations, avoiding potential legal issues.

User: Ron Lev Tabuchov, Team Collaborator

Goal: To maintain an up-to-date GitHub repository for code and documentation.

Benefit: Facilitates seamless collaboration, ensuring that all team members have access to the latest updates and can contribute effectively.

Role: Mohamed Khaled Ahmed Mohamed Elshami, Quality Assurance Engineer

Goal: To develop and implement comprehensive testing strategies for all software components.

Benefit: Ensures that the software is reliable, stable, and meets all technical and user requirements before deployment.

Role: Mohamed Khaled Ahmed Mohamed Elshami, User Experience Designer

Goal: To conduct A/B testing to determine the most effective interface design.

Benefit: Results in an interface that is both intuitive and efficient, significantly improving the end-user experience.

Role: Daria Vladutu, Software Developer

Goal: To establish a consistent, high-standard coding practice including documentation, error handling, and logging.

Benefit: Reduces bugs and improves maintainability of the code, making it easier to manage and update.

7   Control

By integrating Trello for task management, GitHub for version control and code reviews, individual work logs for daily tracking, and regular reviews and feedback sessions via Teams and SharePoint, we can effectively control and manage the project to meet all requirements.

8   Appendices

Server:

http://194.171.191.226:3078/y1d/<STUDENTNUMBER>/lab

Trello Board:

https://trello.com/b/vX1azUlO/dsai-y1d-18

SharePoint Folder:

Y1D - Group 18

Group GitHub:

https://github.com/BredaUniversityADSAI/2023-24d-fai1-adsai-teamwork-t18

Assessment Guidelines 5.0 & 6.0:

https://github.com/BredaUniversityADSAI/2023-24d-fai1-adsai-teamwork-t18/blob/main/Assessment%20Guidelines%20for%20ILO%205.pdf

https://github.com/BredaUniversityADSAI/2023-24d-fai1-adsai-teamwork-t18/blob/main/Assessment%20Guidelines%20for%20ILO%206.0.pdf