

NORWEGIAN UNIVERSITY OF SCIENCE
AND TECHNOLOGY



TDT4259 - APPLIED DATA SCIENCE

Group Assignment

A Data-Driven Approach to Lifestyle Risk Prediction Using Clinical Biomarkers: Smoking and Drinking Analysis

Authors:

Marco Prosperi - 151613
Andrea Richichi - 151790
Tizita Belachew Tamirat - 128549
Gianluigi Vazzoler - 152698

Fall, 2025

Contents

1	Introduction	3
1.1	Background and Context	3
1.2	Problem Definition	3
1.3	Motivation	3
1.4	Team and Roles	3
1.5	Report Outline	3
2	Project Objectives and Scope	3
2.1	Project Goals	3
2.2	Scope and Limitations	3
2.3	Broader Context and Relevance	3
3	Data Strategy and Management	3
3.1	Data Sources and Description	3
3.2	Data Collection and Preprocessing	3
3.3	Data Quality and Cleaning	3
3.4	Data Strategy Framework	3
3.5	Tools and Technologies Used	3
4	Methods and Modeling	3
4.1	Analytical Approach	3
4.2	Modeling Techniques and Algorithms	3
4.3	Model Training and Validation	3
4.4	Interpretability and Explainability	3
4.5	Limitations of Methods	3
5	Analysis and Results	3
5.1	Exploratory Data Analysis	3
5.2	Results Presentation	3
5.3	Interpretation of Findings	3
5.4	Benchmarking and Comparisons	3
5.5	Business or Scientific Implications	3
6	Recommendations and Deployment	3
6.1	Recommendations and Action Plans	3
6.2	Deployment Strategy	4
6.3	Implementation Roadmap	4
6.4	Limitations and Future Work	4
7	Monitoring and Maintenance	4
7.1	Key Performance Indicators (KPIs)	4
7.2	Monitoring Plans and Dashboards	4
7.3	Risk Management and Contingency Plans	4
7.4	Maintenance and Updates	4
7.5	Lessons Learned and Feedback	4
8	Conclusion	4
9	References	4

10 Appendices**4**

1 Introduction

1.1 Background and Context

1.2 Problem Definition

1.3 Motivation

1.4 Team and Roles

1.5 Report Outline

2 Project Objectives and Scope

2.1 Project Goals

2.2 Scope and Limitations

2.3 Broader Context and Relevance

3 Data Strategy and Management

3.1 Data Sources and Description

3.2 Data Collection and Preprocessing

3.3 Data Quality and Cleaning

3.4 Data Strategy Framework

3.5 Tools and Technologies Used

4 Methods and Modeling

4.1 Analytical Approach

4.2 Modeling Techniques and Algorithms

4.3 Model Training and Validation

4.4 Interpretability and Explainability

4.5 Limitations of Methods

5 Analysis and Results

5.1 Exploratory Data Analysis

5.2 Results Presentation

5.3 Interpretation of Findings

5.4 Benchmarking and Comparisons

5.5 Business or Scientific Implications

6 Recommendations and Deployment

6.1 Recommendations and Action Plans

6.2 Deployment Strategy

6.3 Implementation Roadmap

6.4 Limitations and Future Work

7 Monitoring and Maintenance

7.1 Key Performance Indicators (KPIs)

7.2 Monitoring Plans and Dashboards

7.3 Risk Management and Contingency Plans

7.4 Maintenance and Updates

7.5 Lessons Learned and Feedback

8 Conclusion

9 References

10 Appendices