



Data Mining II Project:

Analyzing Data Insights from the IMDb Platform

Authors:

Academic Year 2024/2025

Contents

1	Introduction	2
2	Data Understanding and Preparation	3
3	Outliers	4
3.1	COF	4
3.2	Isolation Forest	4
3.3	ABOD	4
4	Imbalanced Learning	5
4.1	Undersampling	5
4.2	Oversampling	5
4.2.1	SMOTE	5
5	Advanced Classification	6
6	Advanced Regression	6

1 Introduction

2 Data Understanding and Preparation

3 Outliers

3.1 COF

3.2 Isolation Forest

3.3 ABOD

4 Imbalanced Learning

4.1 Undersampling

4.2 Oversampling

4.2.1 SMOTE

5 Advanced Classification

6 Advanced Regression