**Final Report**

**Abdullah Khan 20212022189**

**Jehad Mukahal 20212022240**

**Daniyar Sabitov 20212022029**.

Our project aims to end poverty in all its forms everywhere (SDG 1). We use a dataset containing poverty rates from 5 different countries i.e. USA, Canada, Turkey, Spain and Germany. We employ the Cross-Industry Standard process for data mining (CRISP-DM) to analyze and understand the data which help us find insights to poverty trends and the factors influencing it. Using these results we will find a reliable solution which helps us tackle the poverty issue in OECD countries.

Application Areas:

1. Social Policy Analysis: Understanding the factors influencing poverty rates which results in deploying better policy-making decisions.

2. Economic Development: Analyzing economic indicators that have some relation with poverty trends to facilitate sustainable development in OECD countries.

**Cross Industry Standard Process for Data Mining:**

1. Business Understanding:

- Objective: Analyzing the poverty trends in OECD countries to better understand the impact of economic and social factors on poverty rates.

- Goals: Identify key factors contributing to changes in poverty rates, provide insights for policy recommendations, and predict future trends.

2. Data Understanding:

- Data Sources: Data is taken from different sources that are all cited in our project

- Data Description: The dataset includes poverty rates from 2000 to 2020 for 5 different countries i.e. Turkey Spain, Germany, USA and Canada.

3. Data Preparation:

- Data Cleaning: We use this to prepare data by correcting missing values, inconsistencies and by normalizing the data.

- Feature Selection: Using selective features such as poverty rates or unemployment rates that help us find results for our problem.

4. Modeling:

- Techniques Used:

- Regression Analysis: To predict future poverty rates based on historical data.

- Classification: To predict future poverty rates while using test-train split method to check accuracy and making a new data i.e poverty rate mean as threshold to predict future poverty rates for each country.

- Clustering: To identify groups of countries with similar poverty trends.

5. Evaluation:

- Metrics: Use mean absolute error (MAPE) and R-squared (R²) for regression models; accuracy, precision, and recall for classification models.

- Validation: Cross-validation techniques to ensure model robustness.

6. Deployment:

- Generate comprehensive reports with visualizations to present findings.

- Provide actionable insights for policymakers to address poverty.

Solution:

We use multiple data mining techniques such as classification, regression, and clustering to help identify patterns and relationships in the data, which helps us to understand the underlying causes of poverty trends in different countries. This project will provide valuable insights into poverty trends and contribute to the development of effective policies to combat poverty in OECD countries.

Tasks and Objectives

- Task 1: Analyze historical poverty trends using regression models to understand past changes and make future predictions.

- Objective: Predict future poverty rates so that targeted interventions can be done to prevent country at risk.

- Task 2: Classify using poverty rate as threshold to make future precitions and check model accuracy..

- Objective: Train our model using 80% of our data set and then run it on the remaining 20% of the dataset to check which countries would go above or below the threshold in the future.

- Task 3: Cluster countries based on similar poverty trends and used elbow method to determine clusters.

- Objective: Making groups of countries with similar data to use them for collaborative policy making decisions.

References

1. Zaidi, A. (March 26, 2009). Poverty and Income of Older People in OECD Countries. Available at SSRN: <https://ssrn.com/abstract=1992492> or [http://dx.doi.org/10.2139/ssrn.1992492](https://dx.doi.org/10.2139/ssrn.1992492)
2. Yüksel, H. (2014). General Overview on Poverty: The Sample of OECD Countries. Çankırı Karatekin Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 4(1), 341-358. https://www.researchgate.net/profile/Hasan-Yueksel-6/publication/366217653\_General\_Overview\_on\_Poverty\_The\_Sample\_of\_OECD\_Countries/links/63986d9a484e65005b091397/General-Overview-on-Poverty-The-Sample-of-OECD-Countries.pdf