Universidad Politécnica
de Madrid

Escuela Técnica Superior de Ingenieros Informáticos

Grado Matemáticas e Informática

Fuzzy Countries

In this project, a socioeconomic model for various countries is developed using fuzzy logic. The model is implemented in Ciao Prolog with the RFuzzy library and using Python with Scikit-learn to compare the results with real data and assess the model's credibility. Uflese is used for visualizing the results.

Authors: Javier Comyn, Diego Fogued, Francisco J. González
Professor: Susana Muñoz Hernández

Madrid, 2023/2024

Contents

1 Introduction

1.1 Background and motivation for the study

The motivation for this study comes from the need to better understand and model the complex socioeconomic dynamics of different countries. Traditional economic models often can't handle the uncertainty and vagueness in real-world data. Fuzzy logic theory is well-suited for this task, providing a way to deal with these uncertainties. This study aims to create a more accurate and reliable socioeconomic model, which will be compared with real data to ensure its credibility.

At the beginning, when we started thinking about the project, we were looking for a topic that could be both fascinating and challenging, while also fitting well with the principles of fuzzy logic. We aimed to choose a topic applicable to real life, allowing us to draw conclusions that we might not have realized without applying these tools.

The first thought was about the possibility of focusing on psychological analysis or something related to human mental health because we thought it would be interesting to apply fuzzy logic to this field. However, we quickly realized that this topic was too broad and complex for the scope of our project and also that it would be difficult to find reliable data to work with.

Without giving up on the idea of working with human behavior, we decided to focus on a topic that would allow us to analyze human behavior in a more indirect way. We thought about the possibility of analyzing the relationship between socio-economic and environmental indicators and how these factors can influence the happiness of a country's population. We believe that this topic is relevant and interesting because it allows us to explore the relationship between different aspects of human life and how they can affect people's well-being.

Moreover, this idea of analyzing the happiness of a country's population gives us the opportunity to contrast the results obtained with the World Happiness Report, which is a well-known study that ranks countries based on their happiness levels. This will allow us to validate our results and compare them with those obtained by other researchers, thereby assessing the credibility of this approach.

1.2 Research objectives

The main objective of this research is to develop a socioeconomic model that gives relevant insights into the economic and environmental conditions of different countries, which would not be possible with traditional models and classical logic Additionally, the research seeks to use Uflese for visualizing the outcomes, ensuring that the model's findings are both understandable and useful for further analysis. Ultimately, the goal is to establish a credible model that can provide valuable insights into the socioeconomic conditions of various countries.

As mentioned above, the main objective of this project is to analyze the relationship between socio-economic and environmental indicators and the happiness of a country's population. To achieve this, we will develop a fuzzy logic system with functions and rules to model this relationship and draw conclusions from the available data.

We will use data from reputable sources like the World Happiness Report and the World Bank. Our fuzzy logic system, with functions and rules will process this data to reveal patterns and insights that may not be immediately apparent, and we will find the credibility of our results by comparing the happiness scores we obtain with those in the World Happiness Report.

2 Theorical Framework

2.1 Fuzzy Logic

Fuzzy logic is a form of many-valued logic in which the truth values of variables may be any real number between 0 and 1 both inclusive. It is employed to handle the concept of partial truth, where the truth value may range between completely true and completely false. By contrast, in Boolean logic, the truth values of variables may only be the integer values 0 or 1.

Fuzzy logic has been extended to handle the concept of partial truth, where the truth value may range between completely true and completely false. Furthermore, when linguistic variables are used, these degrees may be managed by specific functions.

3 Methodology

The methodology used in this project can be divided into the following steps:

1. Data Collection: Collecting data from different sources related to socio-economic and environmental indicators.
2. Data Description: Describing the data collected and analyzing its characteristics.
3. Data Preprocessing: Cleaning, transforming, and integrating the data to make it suitable for analysis.
4. Database Design and Development: Designing and developing a database to store the data and integrate it with the fuzzy logic system.
5. Implementation of the Fuzzy System: Developing the fuzzy logic system with functions and rules to model the relationship between the indicators and happiness.
6. Results and Discussion: Presenting and analyzing the results obtained from the fuzzy logic system.

7. Challenges and Solutions: Identifying difficulties encountered during the project and proposing solutions to overcome them.

8. Conclusions and Future Work: Drawing conclusions from the study and suggesting possible future research directions.

4 Database Design and Development

4.1 Data Collection

To gather the data, we used a variety of sources (mainly Kaggle) to obtain information on different socio-economic and environmental indicators for various countries. We analysed which indicators would be most relevant for our study and selected the most reliable and up-to-date datasets available. Furthermore, we ensured that the data was clean and consistent by performing data cleaning and validation procedures.

First of all, we started by collecting data from different sources.

4.2 Data Description

The variables in the database include a mix of socio-economic and environmental indicators, which are:

- GDP per capita
- Renewable energy consumption
- Minimum wage
- Population density
- Economic freedom index
- Mean temperature
- Life expectancy
- Happiness index

4.3 Data Preprocessing

Before integrating the data into the database, we performed several preprocessing steps to clean and transform the data. This included handling missing values, normalizing the data, and

converting categorical variables into numerical values. Additionaly, the different datasets were merged and integrated into a single database, ensuring that the data was consistent and ready for analysis.

4.4 Data Analysis

5 Implementation of the Fuzzy System

6 Results and Discussion

6.1 Querys

7 Challenges and Solutions

7.1 Identified Difficulties

7.2 Overcoming Difficulties

8 Conclusions and Future Work

References

[1] Elgiriyewithana, Nidula, *Global Country Information Dataset 2023.* [Data set]. (2023, 8 julio). Kaggle.
https://www.kaggle.com/datasets/nelgiriyewithana/countries-of-the-world-2023

[2] Hossainds, Belayet, *Renewable Energy world Wide : 1965 2022.* [Data set]. (2023, 3 marzo). Kaggle.
https://www.kaggle.com/datasets/belayethossainds/renewable-energy-world-wide-19652022

[3] Pei Pei Chen, *Minimum wage by country.* [Data set]. (2020, 27 diciembre). Kaggle.
https://www.kaggle.com/datasets/peipeichen/minimum-wage-by-country

[4] My Koryto, *countryinfo.* [Data set]. (2020, 14 abril). Kaggle.
https://www.kaggle.com/datasets/koryto/countryinfo

[5] Fraser Institute, *Economic Freedom of the World.* [Data set]. (2021).
https://www.fraserinstitute.org/economic-freedom/dataset?geozone=world&min-year=2&max-year=0&filter=0&page=dataset&year=2021

[6] Palinatx, *Mean temperature for countries by year 1901-2022.* [Data set]. (2024, 21 marzo). Kaggle.
https://www.kaggle.com/datasets/palinatx/mean-temperature-for-countries-by-year-2014-2022/suggestions?status=pending&yourSuggestions=true

[7] World Health Organization, *Indicators.* [Data set]. (n.d.). World Health Organization.
https://data.who.int/es/indicators

[8] Beach, J., *World Happiness Report 2013-2023* [Data set]. (2023). Kaggle.
https://www.kaggle.com/datasets/joebeachcapital/world-happiness-report-2013-2023

[9] Singh, A. P., *World Happiness Report 2021* [Notebook]. (2021). Kaggle.
https://www.kaggle.com/code/ajaypalsinghlo/world-happiness-report-2021-world/notebook