Exploring Visualization Techniques with World Happiness Dataset

[1]Gitanjali Koli, [2] **Dhanashree Deore**, **[3]Tanmay Deokar**,
[4]Roshani Gaikwad

Department of Information Technology, MET's Institute of Engineering, Adgaon, Nashik

Abstract

Happiness is the current important research issues in psychology and social sciences, which is affecting people's daily lifestyle, work habits and thinking patterns, it also provides guidance for government policy making. However, in the current analysis of happiness, there are many challenges in variable selection and prediction. Due to the large personal differences and differences in determinants of the experience of happiness, this undoubtedly makes the modeling of happiness more difficult. The main purpose of this project is to propose various algorithms to analyses the relationship between people's happiness index and the responses to the problems they have answered using the obtained dataset of the survey, which collects a random number of people's happiness index and to generate predictions in the new dataset that are fairly consistent with actual values.

INTRODUCTION

Predicting the happiness index using learning techniques is an interesting and relevant endeavor in today's data-driven world. Often derived from surveys and assessments of subjective well-being, the happiness index is a multifaceted measure of society's overall satisfaction and life satisfaction. Understanding and predicting the factors influencing happiness can provide valuable insights for policymakers, researchers, and individuals seeking to improve well-being and societal well-being.

Machine learning, a subset of artificial intelligence, offers a powerful set of tools for analyzing patterns and relationships within different data sets. Using advanced algorithms and models, machine learning focuses on uncovering hidden patterns in data and making predictions or classifications based on learned features.

The prediction process generally involves a number of steps, including data preprocessing, feature selection, model selection, training, evaluation, hyper parameter tuning, and interpretation. During model training, selected machine learning models learn to recognize patterns and relationships between features and the happiness index, with the goal of minimizing prediction error.

The ultimate goal is to develop a predictive model that accurately estimates the happiness index based on the provided features. Evaluating the model's performance and iteratively refining the approach makes it possible to make increasingly accurate predictions and shed light on factors that significantly contribute to the overall happiness of a population.

Predicting the happiness index holds enormous potential for helping policymakers formulate informed decisions, promote social reforms, and create strategies to improve societal well-being. By using learning techniques, we can contribute to a deeper understanding of happiness and work to create happier and more fulfilled communities and nations.

LITURATURE SURVEY

•   This paper utilizes machine learning to analyze the features related to the happiness index and to make some predictions. Using the obtained dataset of the survey that collects a random number of Chinese people's happiness index and asks relevant questions, we provide several algorithms to analyze the relationship between people's happiness index and the answers to the problems they have answered. We can make predictions in the new dataset approximately consistent with the actual values by utilizing the results. In our study, we gain some important features like income, education, and health. This paper provides some novel perspectives to improve service for e-governance.[1].

•   In this research paper, the purpose of this paper is to conduct astudy on World happiness report dataset, to classify the most criticalvariables regarding the life happiness score. The strong evidence ofthe identified main features classified from the outcomes of applyingthe supervised machine learning approaches using the NeuralNetwork training model and the OneR models in classifications andfeature selection. The trained model used in predictions revealed theinsights derived from applying the data analysis, where the studyfound out that the GDP per capita is the critical indicator of lifehappiness score as well as the health life expectancy is the secondprimary feature. Findings from study evaluated using differentperformance metrics such as accuracy and confusion matrix toprove the insights gained from the data.[2].

•   This research works with the World Happiness Report 2019 and aims to use machine learning, artificial intelligence, computational strategy. In particular, different machine learning tools such as Google Colab and weka is used in this paper to model the processed historical happiness index report. Using the data of 156 countries from the UN Development Project 2019, this work can identify which factors need to be improved by a particular country to increase the happiness of its citizens. The paper presents supervised machine-learning-based analytical models that can predict the life satisfaction score of any specific country based on the defined parameters, emphasizing the happiest countries and regions based on 2019 happiness report findings. [3].

•   This paper proposes a system for general population sentiment monitoring from a social media stream (Twitter), through comprehensive multi-level filters, and our proposed improved latent Dirichlet allocation (LDA) (Wang et al. in ACM Trans Internet Technol 18(1):1–23, 2017; Wang and Al-Rubaie in Appl Soft Comput 33:250–262, 2015; https://patents.google.com/patent/US20170293597A1/en) method for sentiment classification. Experiments show that our proposed improved LDA for sentiment analysis yields the best results, and also validate our proposed system for national sentiment monitoring in Abu Dhabi using twitter. [4].

In this paper, measuring happiness in quantifiable terms is a global phenomenonlately. United Nation's World Happiness Report (WHR) is one suchmeans to analyses the level of subjective wellbeing that countries acrossthe world are living with. The Happiness Index is framed to set variousparameters on grounds of which a country could be ranked in a list of156 countries. India's rank has come down the list this year (2019) tobe ranked at the 140th position. This clearly indicates India'sdeteriorating position down the years. This paper elaborates theconcept of Happiness Index as a measure and analyses various reasonsfor India to lose its position in the World Happiness Report. The authorappropriately concludes the paper with suitable suggestions. [5]

AIM & OBJECTIVES

•   The primary aim is to conduct a comprehensive analysis of happiness, with a focus on understanding how it impacts people's daily lifestyles, work habits, and thought patterns. This involves delving into the multidisciplinary aspects of happiness in psychology and social sciences.

• The main objective of this project is to build Create predictive models that utilize the dataset to generate accurate predictions of happiness values for new datasets, with the goal of achieving high consistency with actual values.

MOTIVATION

Firstly, happiness is a fundamental aspect of human existence, influencing our daily lives, work habits, and overall well-being. It is an essential component of a fulfilling life, and understanding its complexities can lead to improved quality of life for individuals and communities. The desire to unravel the intricacies of happiness arises from our innate curiosity about what contributes to human contentment and how we can enhance it.

SYSTEM ARCHITECTURE
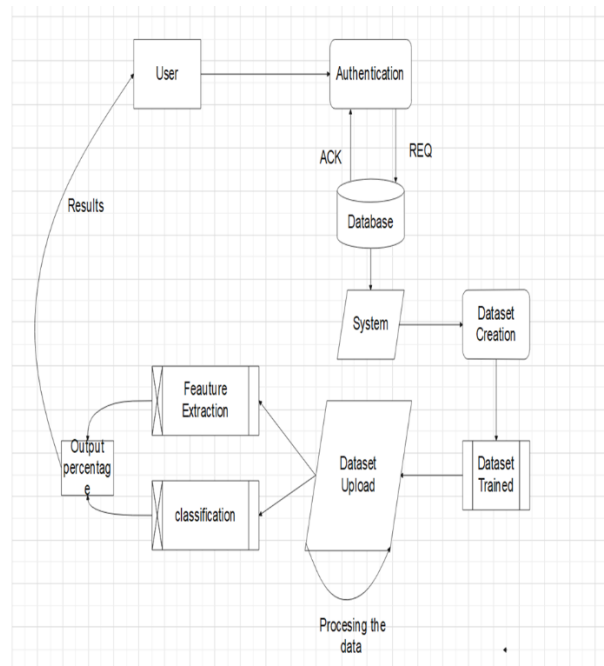
Fig -1: System Architecture Diagram

APPLICATION

• Education and School Counseling: Educational institutions can apply happiness analysis to improve student well-being and academic performance. Health care center.

• Telehealth and Remote Monitoring: In the context of telehealth and remote monitoring, healthcare providers can integrate predictive models to monitor patients' well-being and detect early signs of mental health concerns.

• Human Resources and Talent Management: Human resource departments can leverage predictive models to assess and manage employee happiness and job satisfaction.

• Social Services Allocation: Social service agencies can allocate resources more efficiently by targeting individuals and communities with lower happiness indices for additional support.

FUNCTIONAL&NON-FUNCTIONAL REQUIREMENTS

Functional requirements: may involve Data Collection Module, Data Preprocessing and Cleaning, Variable Selection and Feature Engineering, Algorithm Development, Model Training and Validation and Predictive Modeling and Prediction Generationprocessing and other specific functionality that define what a system is supposed to accomplish.

Nonfunctional Requirements: (NFRs) define system attributes such as security, reliability, performance, maintainability, scalability, and usability. They serve as constraints or restrictions on the design of the system across the different backlogs.

Functional requirements

1.   Registration
2.   User Login
3.   Data Acquisition
4.   Data Preprocessing
5.   Learning Models
6.   Performance evaluation

Security:

1.   User Identification
2.   Login ID
3.   Modification

Performance Requirement:

1.   Response Time
2.   Capacity
3.   User Interface
4.   Maintainability
5.   Availability

METHODOLOGY

The methodology for the project involves a step-by-step process to effectively analyze the relationship between people's happiness index and their responses to survey questions. The functional requirements include the implementation of a registration system where users can sign up to participate in the survey. Upon registration, users will have personalized login credentials for accessing the system. Data acquisition involves collecting a random sample of people's happiness index through the survey, and this data is then preprocessed to handle any missing values, outliers, or noise. The cleaned dataset is used to train various learning models that can capture the complex relationship between happiness and survey responses. The performance of these models is evaluated using appropriate metrics, ensuring the generated predictions align well with the actual values in a new dataset.

In terms of security, user identification is crucial to ensure the integrity of the data. Each user is assigned a unique login ID, and measures are implemented to prevent unauthorized access or modification of the data. The performance requirements focus on response time, capacity, user interface, maintainability, and availability. The system should provide quick response times to user interactions, handle a large capacity of users and data efficiently, offer a user-friendly interface for seamless interactions, be easily maintainable for updates and improvements, and ensure high availability to minimize downtime and disruptions to users participating in the survey. These considerations collectively contribute to the successful implementation and execution of the project**.**

SYSTEM REQUIREMENTS

Software Used:

- Vs code
- Python 3.10
- Windows 11

Hardware Used:
- I3 processor or above
- 150 GB Hard Disk or above
- 4 GB RAM or above

CONCLUSION

In conclusion, this project endeavors to address the complex and multifaceted nature of happiness within the fields of psychology and social sciences. By collecting and analyzing diverse datasets and employing advanced algorithms, it seeks to shed light on the intricate relationship between individuals' happiness indices and their responses to survey questions, despite the challenges posed by personal differences and determinants of happiness. The project's ultimate goal is to not only advance our understanding of happiness but also offer practical applications by providing valuable insights and policy recommendations that can positively impact individuals' daily lives, workplace environments, and government policies, ultimately contributing to the broader well-being and contentment of society.The heart of the project lies in algorithm development, where cutting-edge machine learning algorithms and statistical models are crafted to analyze the intricate relationship between happiness indices and survey responses.

REFERENCES

[1] NARISA ZHAO , HUAN GAO , XIN WEN , AND HUI LI, Received January 12, 2021, accepted January 15, 2021, date of publication January 19, 2021, date of current version January 27, 2021. "Combination of Convolutional Neural Network and Gated Recurrent Unit for Aspect-Based Sentiment Analysis".

[2] Xingyou Wang1, Weijie Jiang2, Zhiyong Luo, Proceedings of COLING 2021, the 26th International Conference on Computational Linguistics: Technical Papers, pages 2428–2437, Osaka, Japan, December 11-17 2021. "Combination of Convolu- tional and Recurrent Neural Network for Sentiment Analysis of Short Texts".

[3] MohammadAL-Smadi, Mahmoud M.Hammad, Sa'ad A.Al-Zboon, SajaALTawal- beh, ErikCambria, " Gated Recurrent Unit with Multilingual Universal Sentence En- coder for Arabic Aspect-Based Sentiment Analysis" Knowledge based Systems is an international and interdisciplinary journal in the field of artificial intelligence, 2021.

[4] Mohsen Ghorbani, Mahdi Bahaghighat , Qin Xin and Figen Ozen, Journal of ¨ Cloud Computing: Advances, Systems and Applications (2020), ConvLSTMConv network: a deep learning approach for sentiment analysis in cloud computing.

[5] G. Rajagopalan, "Data visualization with python libraries" in A Python Data Analyst's Toolkit, Springer, pp. 243-278, 2020.

[6] S. Dixit, M. Chaudhary and N. Sahni, "Network learning approaches to study world happiness", arXiv preprint, 2020.

[7] J. Sachs, G. Schmidt-Traub, C. Kroll, G. Lafortune and G. Fuller, "Sustainable development report 2019" in New York: Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN), pp. 30-31, 2019.

[8] D. Nieciecka, Predicting happiness-comparison of supervised machine learning techniques performance on a multiclass classification problem, 2019.

[9] J. F. Helliwell, H. Huang and S. Wang, "Changing world happiness", World happiness report 2019, vol. 2, pp. 11-46, 2019.

[10] N. Jaques, S. Taylor, E. Nosakhare, A. Sano and R. Picard, "Multi-task learning for predicting health stress and happiness", NIPS Workshop on Machine Learning for Healthcare, 2016.