

Project Report

TOPIC: WORLD HAPPINESS REPORT



TEAM MEMBERS:

GRACIKA JAIN: *21BAI1008*

NIDHI ANN ALEX: *21BPS1009*

SHIVAM KUMAR JHA: *21BPS1333*

TANISHQ JAISWAL: *21BPS1301*

INDEX

1. INTRODUCTION

1.1 Project Overview

1.2 Purpose

2. LITERATURE SURVEY

2.1 Existing problem

2.2 References

2.3 Problem Statement Definition

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

3.2 Ideation & Brainstorming

4. REQUIREMENT ANALYSIS

4.1 Functional requirement

4.2 Non-Functional requirements

5. PROJECT DESIGN

5.1 Data Flow Diagrams & User Stories

5.2 Solution Architecture

6. PROJECT PLANNING & SCHEDULING

6.1 Technical Architecture

6.2 Sprint Planning & Estimation

6.3 Sprint Delivery Schedule

7. CODING & SOLUTIONING (Explain the features added in the project along with code)

7.1 Feature 1

7.2 Feature 2

7.3 Database Schema (if Applicable)

8. PERFORMANCE TESTING

8.1 Performance Metrics

9. RESULTS

9.1 Output Screenshots

10. ADVANTAGES & DISADVANTAGES

11. CONCLUSION

12. FUTURE SCOPE

13. APPENDIX

Source Code

GitHub & Project Demo Link

INTRODUCTION

1.1 Project Overview

The World Happiness Report Analysis project aims to analyze and gain insights from the annual World Happiness Report. This report measures and ranks the happiness of countries based on various factors such as economic, social, and environmental indicators. The project's primary goal is to understand the relationship between happiness and various socio-economic factors using a machine learning model, ultimately extracting valuable insights that can be used in different domains like understanding the relationship between happiness and productivity in the workplace. The World Happiness Report provides a comprehensive analysis of the factors that contribute to happiness, which can be used by businesses to create policies and initiatives that promote employee happiness.

1.2 Purpose

The "World Happiness Report Analysis" project has a central purpose of conducting a thorough examination of the annual World Happiness Reports, aiming to uncover the key drivers of happiness in different countries. Through data analysis and machine learning techniques, the project seeks to extract meaningful insights and identify the socio-economic factors that exert the most significant influence on happiness. These insights can then be applied to various domains, including governance, business, and workplace well-being, with the overarching goal of promoting happiness and well-being. By providing data-driven recommendations and informed decision-making tools, the project intends to contribute to the betterment of societies and organizations, ultimately fostering greater happiness and life satisfaction.

LITERATURE SURVEY

2.1 Existing problem

The existing problem addressed by the "World Happiness Report Analysis" project revolves around the challenge of understanding the complex relationship between happiness and the multitude of socio-economic factors that impact it. While the World Happiness Report provides valuable data and rankings, businesses and governments often face difficulties in deciphering how to leverage this information effectively. Specifically, there is a need to bridge the gap between the insights provided by the report and their practical application in areas like the workplace. As such, the project seeks to address this problem by utilizing data analytics and machine learning to provide actionable insights that can be utilized to enhance happiness and well-being, particularly in the context of employee productivity and satisfaction.

2.2 References

Analysis of World Happiness Report Dataset Using Machine Learning Approaches - <https://www.i-csrs.org/Volumes/ijasca/2022.1.2.pdf>

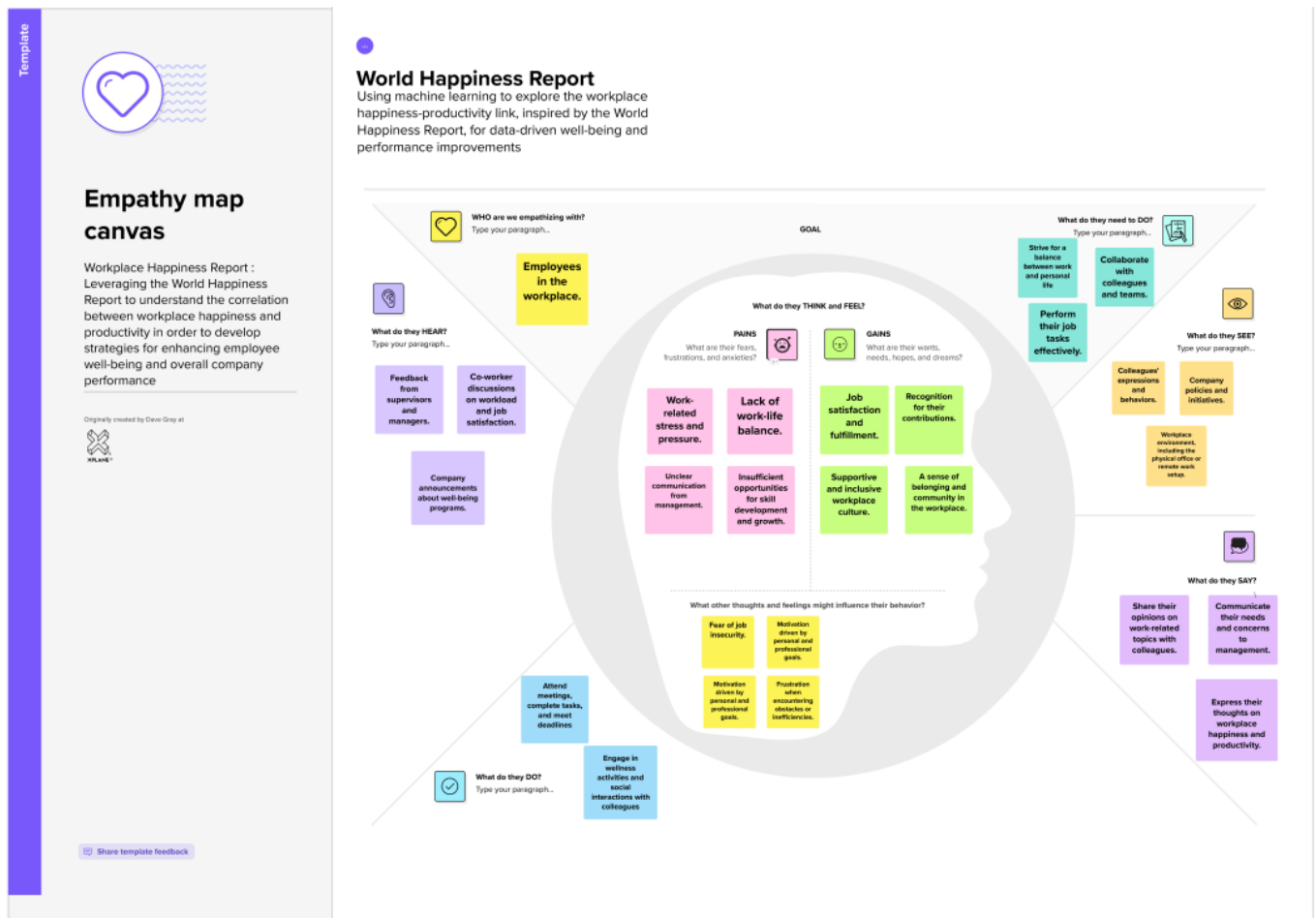
2.3 Problem Statement Definition

Understanding the correlation between workplace happiness and productivity in order to develop strategies for enhancing employee well-being and overall company performance.

IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

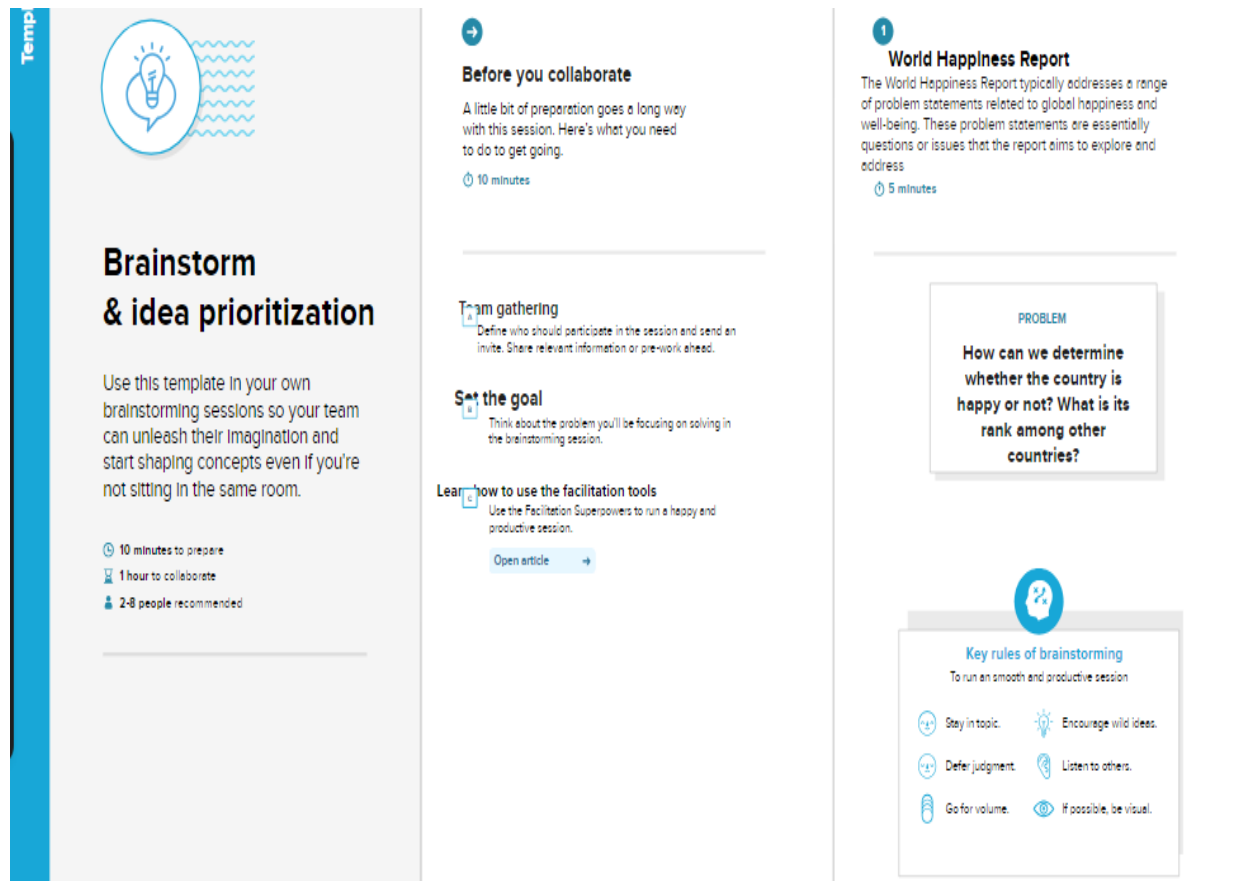
Centered around the World Happiness Report, our project employs an empathy map to delve into the thoughts, emotions, and actions of employees within the workplace. This exploration seeks to establish a profound understanding of their experiences and needs. By empathizing with their aspirations and challenges, we aim to develop data-driven strategies that enhance workplace happiness, which, in turn, translates to heightened productivity and contributes to the overall success of the organization. This approach harnesses emotional intelligence and data insights from the report to craft policies and initiatives that foster a happier, more productive work environment, ultimately benefiting both employees and the company as a whole



3.2 Ideation & Brainstorming

Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich amount of creative solutions.

Step-1: Team Gathering, Collaboration and Select the Problem Statement



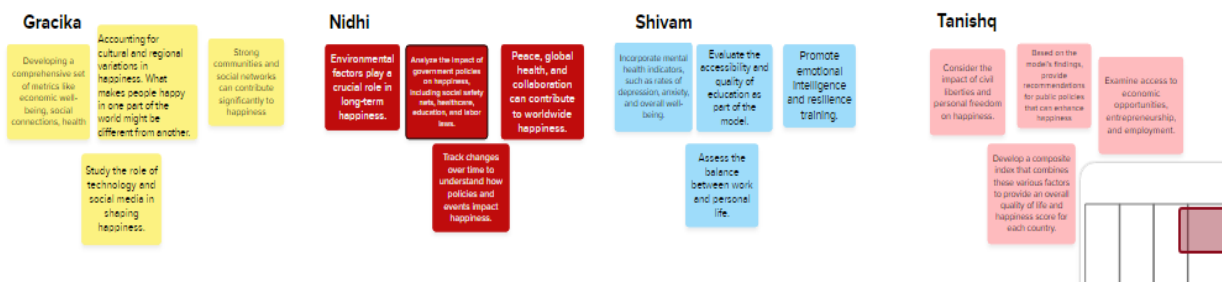
Step-2: Brainstorm, Idea Listing and Grouping

Brainstorm

Writing down our ideas that came to our mind while discussing our problem statement.

10 minutes

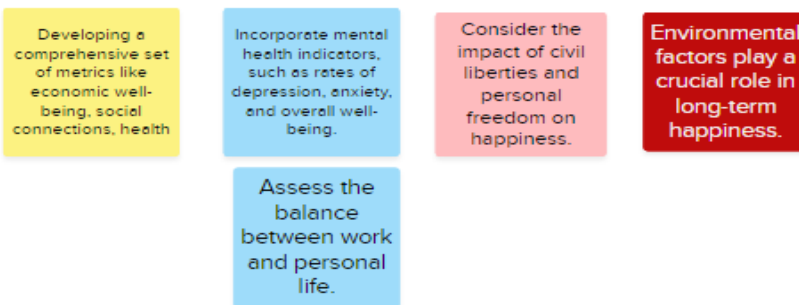
and hit the pencil [switch to sketch] icon to start drawing!



Group ideas

Alternate in expressing our thoughts, grouping akin or interconnected notes as we proceed. After consolidating all the adhesive notes into clusters, assigned a concise label resembling a sentence to each cluster.

🕒 20 minutes



Step-3: Idea Prioritization

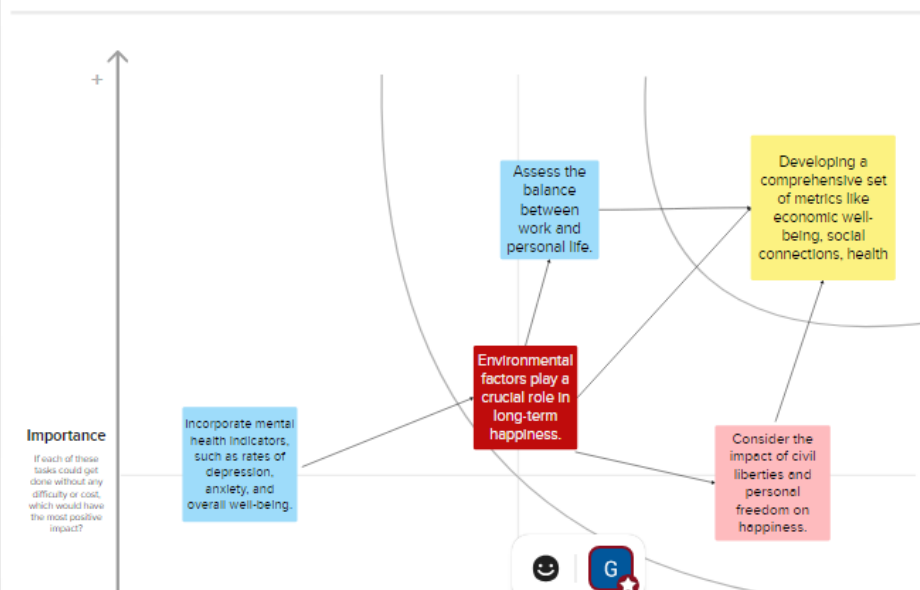
4

Prioritize

It's crucial for our team to have a shared understanding of what holds significance in the future. Using this grid we can assess the importance and feasibility of your ideas.

🕒 20 minutes

TIP
Participants can use their cursors to point at where sticky notes should go on the grid. The facilitator can confirm the spot by using the laser pointer holding the **H** key on the keyboard.



REQUIREMENT ANALYSIS

4.1 Functional requirement

1. Data Collection and Preprocessing:

- The code starts by importing necessary libraries.
- It extracts a zip file and reads the data for the year 2015.
- The dataset is checked for missing values, and as there are no null values, no further preprocessing is needed.
- Univariate analysis is performed by visualizing the 'Happiness Score' using a kernel density plot.
- The code counts countries based on regions and visualizes it.
- A correlational matrix is created, and a heatmap is used to visualize the correlation between different features.
- A pairplot is created for selected columns to visualize relationships.

2. Data for 2016:

- Similar steps are performed on the dataset for the year 2016.

3. Label Encoding and Outlier Removal:

- Label encoding is applied to 'Region' and 'Country' columns for modeling.
- Outliers are removed from the dataset using the Z-score method.

4. Machine Learning Models:

- The data is split into training and testing sets.
- Four machine learning models are implemented:
 - Decision Tree Regressor
 - Random Forest Regressor
 - XGBoost Regressor
 - Linear Regression

5. Model Evaluation:

- For each model, performance metrics like R-squared score and mean squared error are calculated on both the training and testing sets.

6. Hyperparameter Tuning:

- Hyperparameter tuning is performed for the Random Forest model using GridSearchCV.
- The best hyperparameters and model are printed.

7. Validation Method:

- Cross-validation scores and root mean squared error (RMSE) scores are calculated for the Random Forest model.

4.2 Non-Functional requirements

Certainly, here is an explanation of the non-functional requirements point-wise:

1. Performance: This aspect concerns the system's ability to perform efficiently. It includes factors like response time, throughput, and resource utilization. In the context of the code provided, it's important that the machine learning models execute within a reasonable time and can handle large datasets without excessive resource consumption.

2. Data Accuracy: Non-functional requirements necessitate that the data used in the system is accurate. In the code, this implies that the input data is reliable and free from errors. Garbage in, garbage out - accurate data is essential for meaningful results.

3. Scalability: Scalability is about the system's capability to handle growing amounts of data or increasing numbers of users. The code should be designed in a way that it can easily accommodate larger datasets or be extended to include more features or years' worth of data.

4. Security: Security is crucial for protecting data and preventing unauthorized access. In the context of the code, this means that measures should be taken to secure sensitive data, especially if it includes personal information. It's essential to follow best practices for data security.

5. Ethical Considerations: This is a critical non-functional requirement, particularly when working with sensitive data like happiness scores. Ethical considerations entail ensuring that data handling and analysis methods are fair, transparent, and respect the privacy and consent of individuals whose data is being used.

6. User-Friendly Interface: The system should have a user-friendly interface that is intuitive and easy to navigate. In this code, this may refer to well-commented and organized code that is easy for other data scientists or analysts to understand. It could also apply to any user interfaces or dashboards used for data visualization and reporting.

PROJECT DESIGN

5.1 Data Flow Diagrams & User Stories

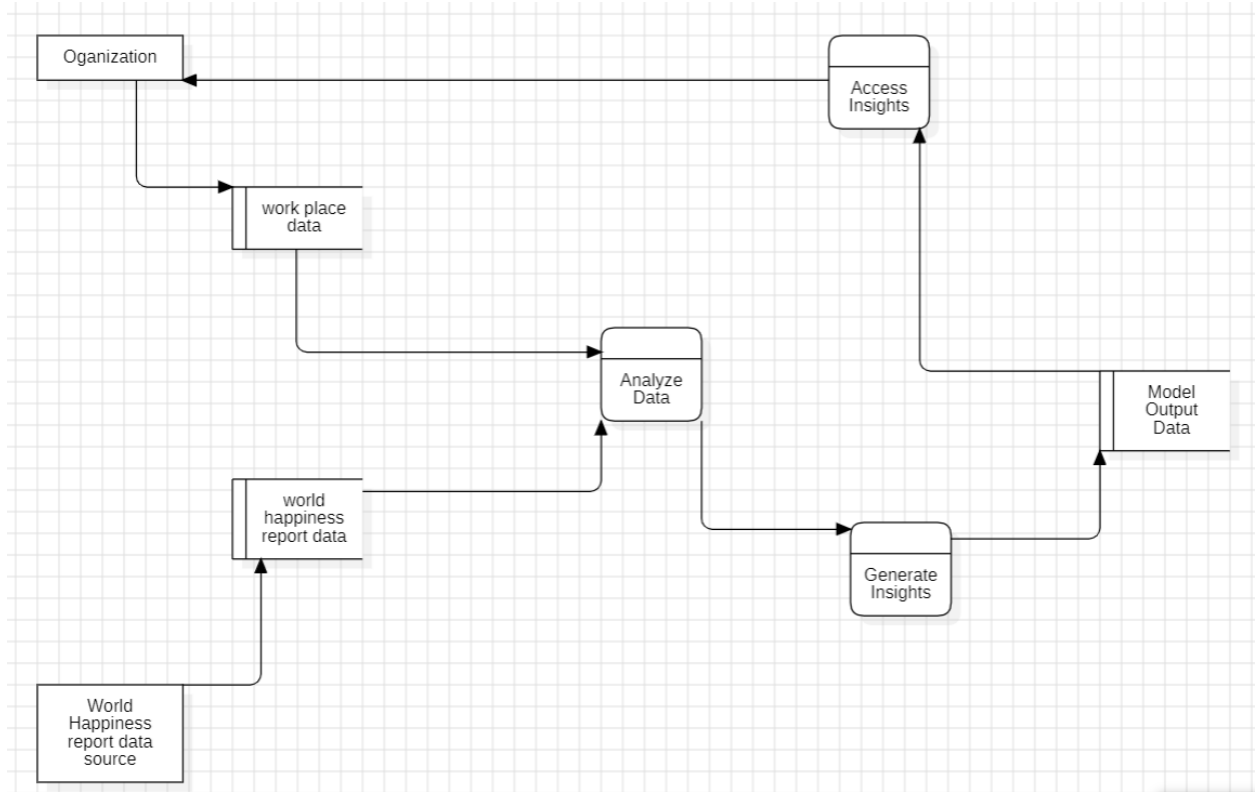
Data Flow Diagrams:

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

Data Flow Diagram for World Happiness Report

1. Data flows from Organizations to Workplace Data Store and from World Happiness Report Data Source to Happiness Data Store
2. Data flows from Workplace Data Store and Happiness Data Store to the process analyze data then the analyzed data goes to the process to generate insights.
3. Modified data now flows to Model Output Data Store

When organizations call the process access data the data moves from model output data store to organizations



User Stories :

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
HR Manager	Registration	USN 1	As an HR Manager, I want to register for the platform.	The registration process should be straightforward, collecting necessary information and creating a secure account.	High	Sprint 1
	Subscription	USN 2	As an HR Manager, I want to subscribe to the platform.	I should be able to choose a subscription plan, provide payment details, and receive confirmation of my subscription.	High	Sprint 1
	Dashboard	USN 3	As an HR Manager, I want to access my dashboard.	I should be able to access options like uploading employee data, reports, suggested policies, etc.	Medium	Sprint 2
	Uploading Employee data	USN 4	As an HR Manager, I want to upload employee data.	The platform should allow me to upload employee information, including names, positions, and relevant metrics securely and in bulk.	High	Sprint 1
	Happiness Reports	USN 5	As an HR Manager, I want to receive automated reports on employee well-being trends.	I should be able to set up automated reports to be sent to my email on a weekly	High	Sprint 1

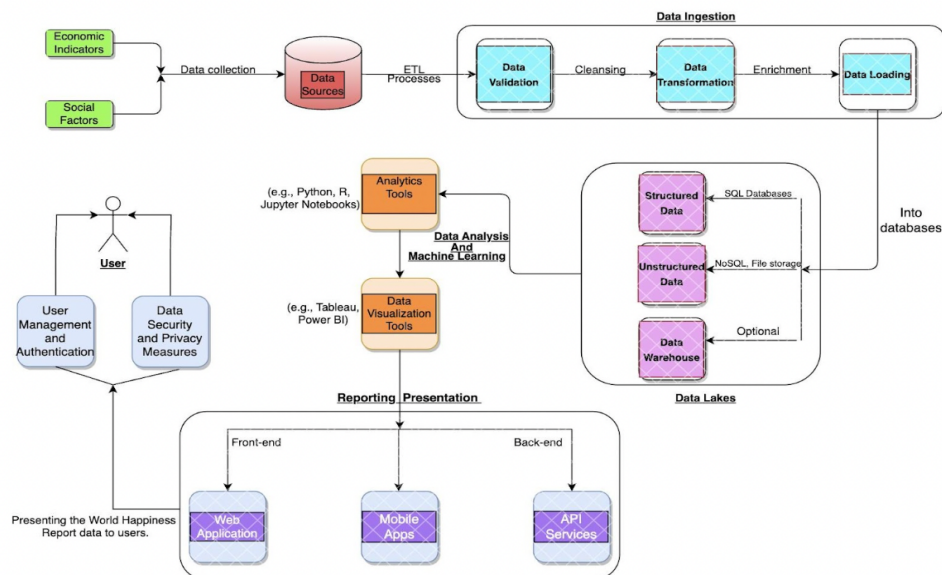
				basis. The reports should include trends, anomalies, and policy recommendations.		
	Recommendation	USN 6	As an HR Manager, I want to generate customized policy recommendations for improving employee well-being.	I should be able to input specific company goals and constraints, and the platform should provide tailored recommendations for policy adjustments.	Medium	Sprint 2
IT Administrator	IT Management and System Maintenance	USN 7	As an IT Administrator, I want to ensure the secure and seamless integration of workplace data into the platform.	I should be able to set up data integration pipelines, validate data accuracy, and schedule regular updates. The system should support various data formats and ensure data encryption in transit.	High	Sprint 1
		USN 8	As an IT Administrator, I want to manage user access and permissions.	The platform should adhere to security best practices, including encryption, access controls, and regular security audits. I should receive alerts for any security breaches.	High	Sprint 1
		USN 9	As an IT Administrator, I want to ensure system performance and reliability.	I should monitor system performance, address any performance issues promptly, and apply software	High	Sprint 1

				updates and patches to maintain system reliability.		
			As an IT Administrator, I want to set up backup and recovery procedures.	I should implement data backup procedures and test data restoration. The platform should have robust backup and recovery capabilities to safeguard data.	Medium	Sprint 1
			As an IT Administrator, I want to set up monitoring and alerts.	I should configure monitoring tools to track system performance, detect anomalies, and receive real-time alerts for system failures, security breaches, or performance issues.	High	Sprint 1

5.2 Solution Architecture

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

- Find the best tech solution to solve existing business problems.
- Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.
- Define features, development phases, and solution requirements.
- Provide specifications according to which the solution is defined, managed, and delivered.



Architecture and data flow of the World Happiness Report

This diagram provides an overview of the key components in the World Happiness Report system:

- **Data Sources:** This is where data from various surveys, economic indicators, and social factors are collected.

- Data Ingestion: The ETL (Extract, Transform, Load) processes for data validation, cleansing, transformation, and loading into databases.
- Data Storage: Structured and unstructured data is stored in databases and data lakes for analysis.
- Data Analysis & Machine Learning: This layer includes tools and frameworks for performing data analysis and machine learning on the stored data.
- Reporting and Presentation: Web applications, APIs, and mobile apps for presenting the World Happiness Report data to users.
- User Access and Authentication: User management and authentication to control access to the data.
- Security and Compliance: Ensures data security and compliance with privacy regulations.

PROJECT PLANNING & SCHEDULING

6.1 Technical Architecture

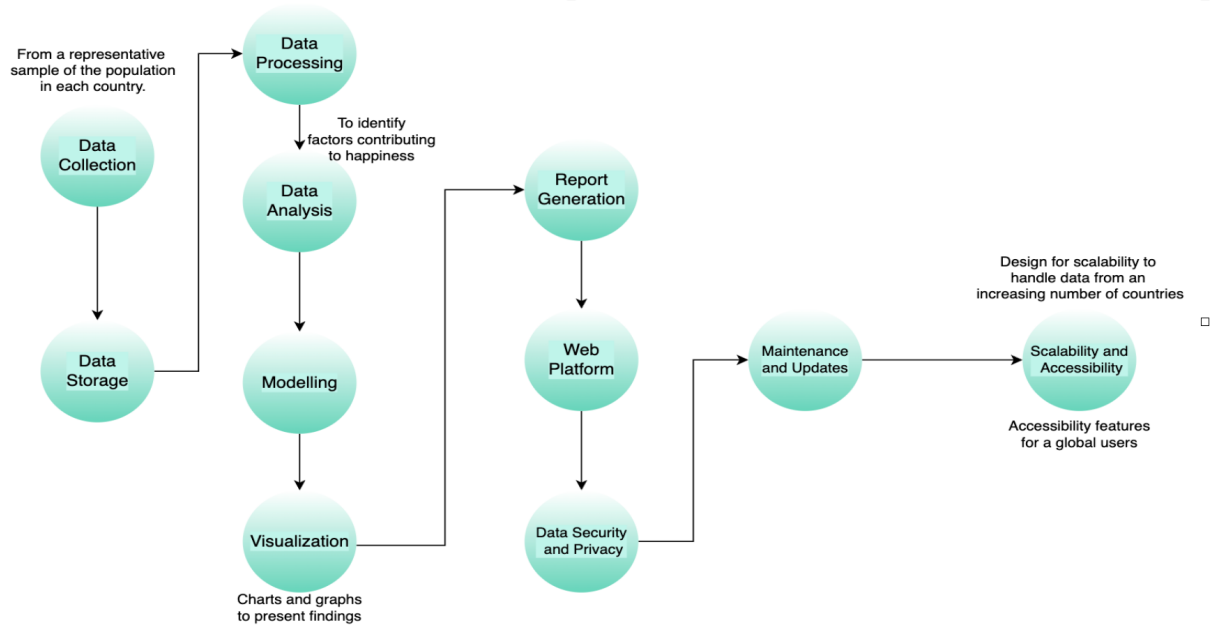


Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	To build the front end of the product through which the users will interact with the system.	HTML, CSS, JavaScript / React Js etc.
2.	Application Logic-1	The process involves making predictions in the background using either a cloud or local system, and allowing users to interact with the predictions through a web interface.	Python/ flask
5.	Database	Data Type, Configurations etc.	Kaggle
6.	Cloud Database	Database Service on Cloud	Flask
7.	File Storage	File storage requirements	Local device
8.	External API-1	To fetch the data from kaggle	kaggle API
10.	Machine Learning Model	Purpose of Machine Learning Model	Regression and Classification Models
11.	Infrastructure (Server / Cloud)	Local server configuration and to deploy an application to the cloud	Flask, Google

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	List the open-source frameworks used	Flask
2.	Security Implementations	To incorporate security measures to protect the system from unwanted threats.	Https
3.	Scalable Architecture	To ensure the future scaling of the System when the traffic or demand is more.	google cloud
4.	Availability	To ensure the availability of the system, it's crucial to use a reliable and redundant infrastructure.	google cloud
5.	Performance	Continuously fine-tune the model to adapt to changing market conditions and improve prediction accuracy.	Metrics

6.2 Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint -1	ML Model	USN-1	Enhance model, for world happiness report	9	High	2
Sprint -1	ML Model	USN-2	Real time data use	5	Medium	1
Sprint -2		USN-3	User authentication	4	Medium	2

Sprint -3	Using different models for report generation	USN-4	Model deployment	12	High	3
Sprint -4	HTML, CSS, javascript	USN-5	Making the UI for web	8	High	2

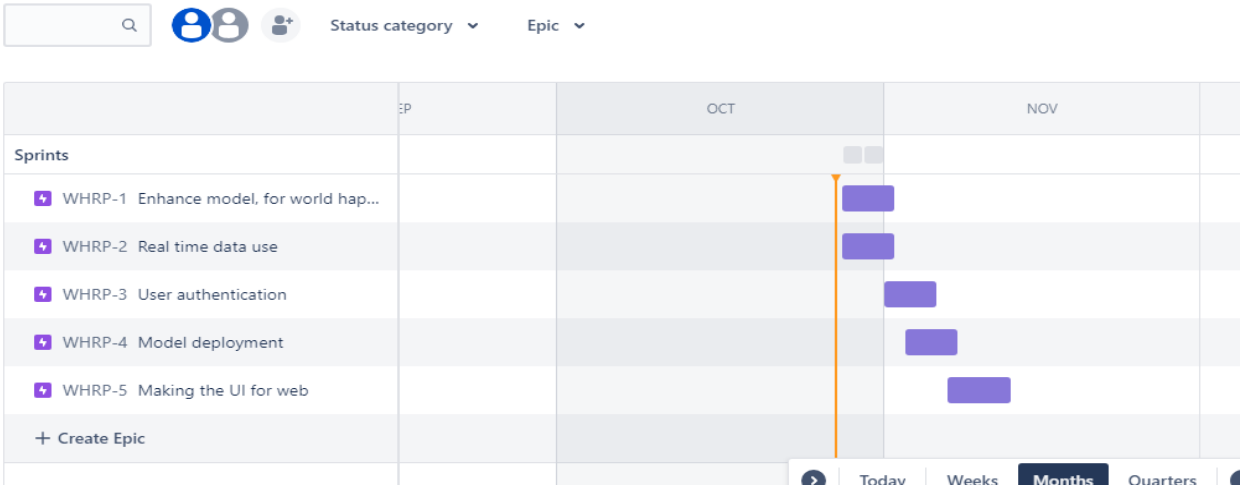
6.3 Sprint Delivery Schedule

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)
Sprint 1	14	4 Days	28 Oct 2023	31 Oct 2023
Sprint 2	4	2 Days	01 Nov 2023	02 Nov 2023
Sprint 3	12	4 Days	03 Nov 2023	06 Nov 2023
Sprint 4	8	4 Days	07 Nov 2023	12 Nov 2023

Projects / World Happiness Report Project

Timeline

[Give feedback](#) [Share](#) [Exp](#)



CODING AND SOLUTION

7.1 Feature 1: Country

It just tells the name of the country which is further ranked on the basis of happiness.

7.2 Feature 2: Region

It is a classification feature dividing the countries on their location.

7.3 Feature 3: Happiness Rank

It is a value telling the position of a country compared to all countries. The basic way of labeling it is higher the happiness score prior to the happiness rank.

7.4 Feature 4: Happiness Score

It is the value that is calculated using different regression models.

7.5 Feature 5: Standard Error

It just signifies the error encountered while calculating the happiness score.

7.6 Feature 6: Economy (GDP per capita)

The gross domestic product (GDP) per capita, which is a measure of a country's economic output per person.

7.7 Feature 7: Family

Just the average number of family members in a house.

7.8 Feature 8: Health(Life Expectancy)

The average number of years of good health an individual can expect to live.

7.9 Feature 9: Freedom

The degree to which people have the freedom to make choices about their lives and live in accordance with their values.

7.10 Feature 10: Trust(Government Corruption)

The level of perceived corruption in government and business.

7.11 Feature 11: Generosity

This measures charitable giving and volunteerism within a country's population.

7.12 Feature 12: Dystopia Residual

It represents a low point on the happiness scale and serves as a reference point for comparing countries' happiness scores.

```
df=pd.read_csv('/content/destination_folder/2015.csv')
df.head()
```

	Country	Region	Happiness Rank	Happiness Score	Standard Error	Economy (GDP per Capita)	Family	Health (Life Expectancy)	Freedom	Trust (Government Corruption)	Generosity	Dystopia Residual
0	Switzerland	Western Europe	1	7.587	0.03411	1.39651	1.34951	0.94143	0.66557	0.41978	0.29678	2.51738
1	Iceland	Western Europe	2	7.561	0.04884	1.30232	1.40223	0.94784	0.62877	0.14145	0.43630	2.70201
2	Denmark	Western Europe	3	7.527	0.03328	1.32548	1.36058	0.87464	0.64938	0.48357	0.34139	2.49204
3	Norway	Western Europe	4	7.522	0.03880	1.45900	1.33095	0.88521	0.66973	0.36503	0.34699	2.46531
4	Canada	North America	5	7.427	0.03553	1.32629	1.32261	0.90563	0.63297	0.32957	0.45811	2.45176

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 158 entries, 0 to 157
Data columns (total 12 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                -
0   Country                               158 non-null    object
1   Region                               158 non-null    object
2   Happiness Rank                       158 non-null    int64
3   Happiness Score                     158 non-null    float64
4   Standard Error                      158 non-null    float64
5   Economy (GDP per Capita)            158 non-null    float64
6   Family                             158 non-null    float64
7   Health (Life Expectancy)            158 non-null    float64
8   Freedom                             158 non-null    float64
9   Trust (Government Corruption)        158 non-null    float64
10  Generosity                          158 non-null    float64
11  Dystopia Residual                    158 non-null    float64
dtypes: float64(9), int64(1), object(2)
memory usage: 14.9+ KB
```

PERFORMANCE TESTING

The best model chosen for the report is RANDOM FOREST MODEL, and the metrics are MAE, R2 SCORE, MSE, RMSE followed by hyperparameter testing and validation method.

Mean Absolute error MAE

```
mean_absolute_error(y_test,y_test_pred1)
```

```
0.05014958333333381
```

R2 Score

```
r2_score(y_test,y_test_pred1)*100
```

```
99.51774792463739
```

Mean squared error MSE

```
mse=mean_squared_error(y_test,y_test_pred1)  
print(mse)
```

```
0.006419666879166671
```

Root Mean Squared Error RMSE

```
rmse = sqrt(mse)  
print(rmse)
```

```
0.08012282370939426
```

Hyperparameter Tuning

```
import numpy as np
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import GridSearchCV
from sklearn.metrics import make_scorer, mean_squared_error
model = RandomForestRegressor(random_state=42)
param_grid = {
    'n_estimators': [50, 100, 200],
    'max_depth': [None, 10, 20, 30],
    'min_samples_split': [2, 5, 10],
    'min_samples_leaf': [1, 2, 4],
    'max_features': ['auto', 'sqrt', 'log2']
}
scorer = make_scorer(mean_squared_error, greater_is_better=False)
grid_search = GridSearchCV(model, param_grid, scoring=scorer, cv=5)
grid_search.fit(x, y)
best_params = grid_search.best_params_
best_model = grid_search.best_estimator_

print("Best Hyperparameters:", best_params)
print("Best Model:", best_model)
```

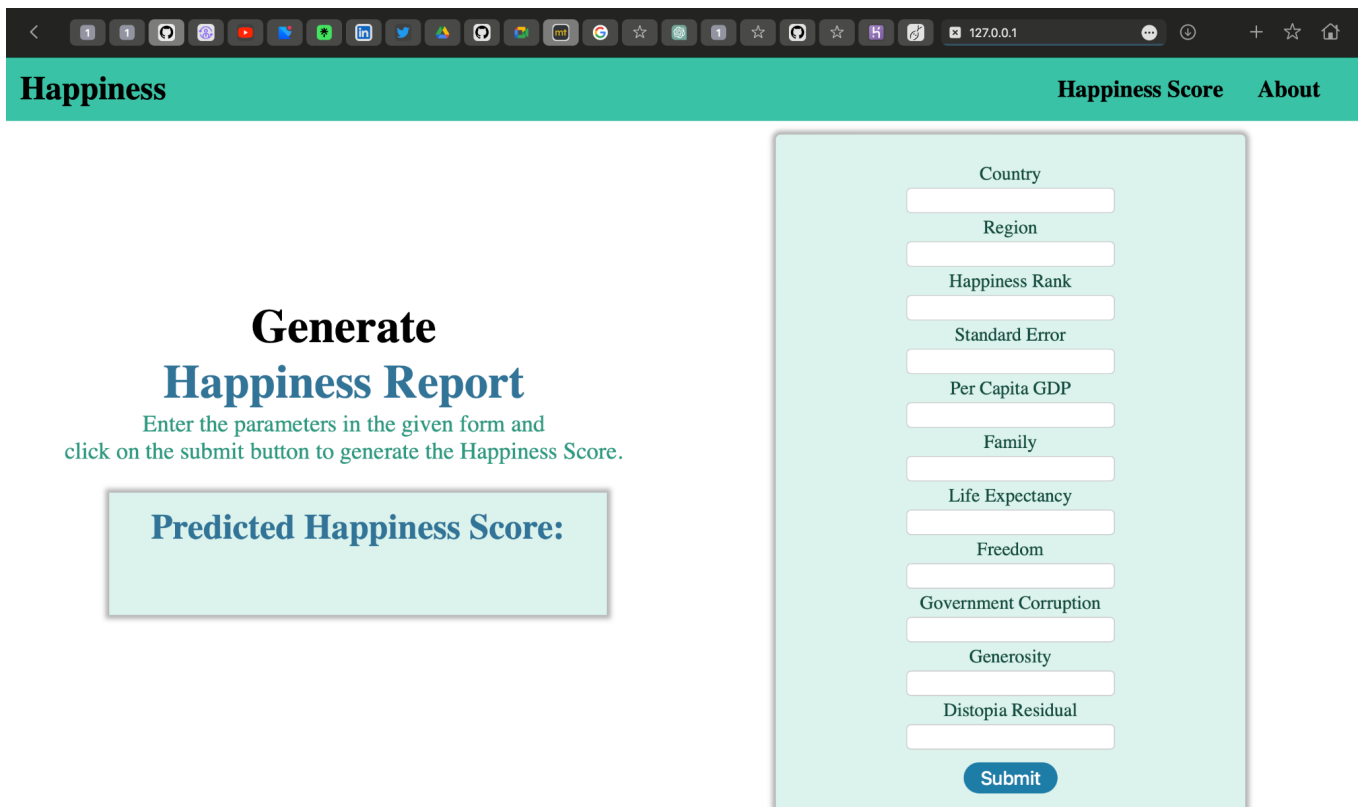
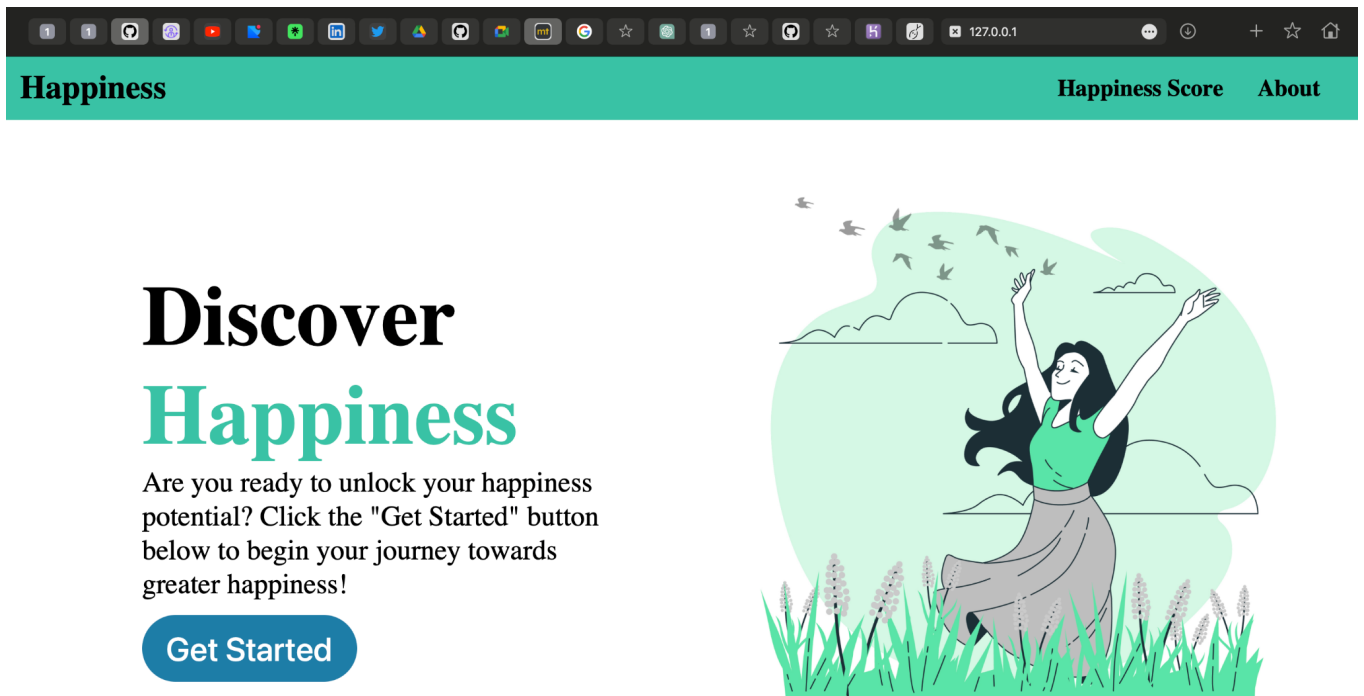
Best Hyperparameters: {'max_depth': 10, 'max_features': 'auto', 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 100}
Best Model: RandomForestRegressor(max_depth=10, max_features='auto', random_state=42)

Validation method

```
[ ] from sklearn.model_selection import cross_val_score
model = RandomForestRegressor(random_state=42)
scores = cross_val_score(model, x, y, cv=5, scoring='neg_mean_squared_error')
rmse_scores = np.sqrt(-scores)
print(rmse_scores)
```

[0.71985686 0.30197233 0.28300866 0.21729802 0.72209794]

RESULTS



Happiness

Happiness ScoreAbout

Generate
Happiness Report

Enter the parameters in the given form and
click on the submit button to generate the Happiness Score.

Predicted Happiness Score:

Country

1

Region

121

Happiness Rank

12

Standard Error

122

Per Capita GDP

12

Family

122

Life Expectancy

12

Freedom

1

Government Corruption

122

Generosity

112

Distopia Residual

11

Submit

Happiness

Happiness ScoreAbout

Generate
Happiness Report

Enter the parameters in the given form and
click on the submit button to generate the Happiness Score.

Predicted Happiness Score:
7.361140000000005

Country

Region

Happiness Rank

Standard Error

Per Capita GDP

Family

Life Expectancy

Freedom

Government Corruption

Generosity

Distopia Residual

Submit

About Us

Welcome to the Happiness Report Generator project. This project aims to predict happiness scores based on various parameters and features. It is designed to provide insights into the well-being of different countries and regions. Our team is dedicated to providing accurate and meaningful happiness scores for analysis and research. If you have any questions or feedback, please feel free to contact us.

Our Team

Gracika Jain
Nidhi Ann Alex
Tanishq Jaiswal
Shivam Kumar Jha

ADVANTAGES & DISADVANTAGES

Advantages:

Comprehensive Assessment: The World Happiness Report utilizes various data sources, including surveys, economic indicators, social and cultural data, health and education statistics, and environmental factors. This comprehensive approach provides a holistic perspective on societal well-being.

Insights for Businesses: The report offers valuable insights for businesses by exploring the connection between workplace happiness and productivity. This information can be used to develop strategies and initiatives aimed at enhancing employee happiness and well-being, leading to improved job satisfaction and overall company performance.

Data Processing Efficiency: AI and machine learning techniques play a crucial role in processing and analyzing the vast amount of data collected from multiple sources. These algorithms automate data preprocessing tasks such as cleaning, normalization, outlier detection, and imputation of missing values, saving time and enhancing data accuracy.

Predictive Modeling: The project employs predictive models powered by AI and machine learning to forecast future happiness trends. These models provide valuable insights into how various factors impact happiness over time, enabling policymakers and organizations to anticipate societal well-being shifts and make informed decisions.

Disadvantages:

Data Privacy Concerns: The project deals with personal survey data, which raises ethical considerations and data privacy concerns. It is crucial to ensure that privacy and data protection measures are in place to address these concerns and protect individuals' sensitive information.

Accuracy and Limitations of Predictive Models: While predictive models offer the potential for foresight, their accuracy and limitations should be critically evaluated. Overreliance on predictive models without considering their limitations may lead to inaccurate outcomes and decisions. It is essential to interpret the results of predictive modeling with caution and consider other contextual factors

CONCLUSION

The World Happiness Report project is a valuable tool for understanding and improving societal well-being. The project's comprehensive assessment of happiness levels across countries and regions provides valuable insights for businesses, policymakers, and organizations. By exploring the connection between workplace happiness and productivity, the report offers actionable information to enhance employee well-being and job satisfaction.

One of the most significant advantages of the World Happiness Report project is its utilization of AI and machine learning techniques in data processing. These technologies automate tasks such as data cleaning, normalization, outlier detection, and imputation of missing values. This not only saves time but also enhances data accuracy, ensuring reliable and robust analysis. Additionally, predictive models powered by AI and machine learning enable the project to forecast future happiness trends. This foresight allows policymakers and organizations to anticipate shifts in societal well-being and make informed decisions accordingly.

The project report also highlights the role of AI and machine learning in creating a web application for generating happiness reports based on user input. This application allows users to input their data and generate customized happiness reports, providing a personalized perspective on their well-being.

However, there are also some disadvantages to consider. One major concern is data privacy. The project deals with personal survey data, which raises ethical considerations and privacy concerns. It is crucial to prioritize data protection measures to ensure the confidentiality and security of individuals' sensitive information.

Furthermore, while predictive models offer valuable insights, their accuracy and limitations should be carefully evaluated. Overreliance on these models without considering their constraints may lead to inaccurate outcomes and decisions. It is essential to interpret the results of predictive modeling with caution, taking into account other contextual factors that may influence happiness levels.

Overall, the World Happiness Report project, with its integration of AI and machine learning, provides a valuable tool for understanding and improving societal well-being. The insights derived from this project have the potential to influence policymaking, shape future initiatives, and prioritize happiness and well-being on a global scale. However, it is essential to address data privacy concerns and carefully evaluate the accuracy and limitations of predictive models. By considering these factors, the project can continue to provide valuable insights into happiness and well-being for years to come.

Future Scope

The future scope of the World Happiness Report project is promising. The project has already made significant contributions to the understanding and improvement of societal well-being, but there are several areas in which it could continue to evolve and expand.

One potential area of future growth is the inclusion of more diverse data sources. While the project currently utilizes a variety of data sources, including survey data, social media data, and economic indicators, there is potential to incorporate additional sources such as environmental data, health data, and education data. By including a broader range of data, the project could provide a more comprehensive understanding of societal well-being and the factors that contribute to it.

Another potential area for growth is the expansion of the project's predictive modeling capabilities. While the project currently employs predictive models to forecast future happiness trends, there is potential to develop more sophisticated models that can identify causal relationships between variables and provide more nuanced insights. These models could be used to identify specific interventions that could improve well-being in different contexts.

Moreover, the project could expand its focus beyond national-level analysis to include regional and local-level analysis. This would allow policymakers and organizations to identify specific areas where interventions are needed and tailor their efforts accordingly. Additionally, it could provide insights into how different regions and communities experience happiness differently.

The integration of AI and machine learning in the project could also continue to evolve. As these technologies continue to advance, there is potential to develop more sophisticated algorithms that can process and analyze data more efficiently and accurately. Additionally, there is potential to develop AI-powered chatbots or other interactive tools that can provide personalized insights into well-being based on individual input.

Finally, the project could continue to prioritize ethical considerations and data privacy. As personal data continues to become more valuable and vulnerable, it is crucial to maintain strict privacy protections and ensure that individuals' sensitive information is not compromised.

APPENDIX

Source Code :  `world_happiness_report.ipynb` (contains ipynb file)

GitHub :

<https://github.com/smartinternz02/SI-GuidedProject-599471-1697529838>

(Contains all design thinking documents + Source code + Project Doc)

Project Demo Link : https://www.youtube.com/watch?v=naDH_AJDs7M