UNIVERSITI MALAYA
Faculty of Computer Science and Information Technology

15 LIFE ON LAND

EXPLORING AGRI-FOOD

Greenhouse Gas Emmision

Group 3

CHENG YEE ERN (22004791) | QUAH JUN CHUAN (22004851) | TER ZHEN HUANG (22004736)

01 INTRODUCTION

PROBLEM STATEMENT

- **Agrifood systems contribute approximately one-third (31%) to total anthropogenic greenhouse gas emissions. These emissions are generated within the farm gate, through crop and livestock production; through land-use change; and in pre- and post-production processes. (FAO, 2021)**

- **These agri-food activities, principally through emissions of greenhouse gasses, have unequivocally caused global surface temperature increase from 1850–1900 to 2010–2019 is 0.8°C to 1.3°C. (NOAA, 2023)**

OBJECTIVES

01 **To identify significant activities contributing to greenhouse gas emissions in the agri-food sector affecting rice production and surface temperature change.**

02 **To classify the surface temperature change of Asian countries due to the emission of greenhouse gases from agri-food activities**

03 **To forecast the future trend of agri-food (rice) production among each Asian countries**

Source : Food and Agriculture Organization of the United Nations (FAO), National Oceanic and Atmospheric Administration(NOAA)

02a | DATA COLLECTION

DATA SOURCE

ABOUT DATA

1. Emission of **greenhouse gases** (CH4, N2O and CO2) from agricultural activities for **Asian countries** (1991-2021) - FAO
2. **Rice production quantity** for **Asian countries** (1991-2021) - FAO
3. Annual **surface temperature index change** for **all countries** (1961-2023) - IMF

02b | DATA WRANGLING

	Domain	Area	Element	Item	Year	Unit	Value
0	Emissions totals	Afghanistan	Emissions (N2O)	Crop Residues	1991	kt	0.7906
1	Emissions totals	Afghanistan	Emissions (N2O)	Crop Residues	1992	kt	0.7416
2	Emissions totals	Afghanistan	Emissions (N2O)	Crop Residues	1993	kt	0.8710
3	Emissions totals	Afghanistan	Emissions (N2O)	Crop Residues	1994	kt	0.9134
4	Emissions totals	Afghanistan	Emissions (N2O)	Crop Residues	1995	kt	0.9201

(82363, 7)

	Country	Year	Emissions (CH4)_Agrifood Systems Waste Disposal	Emissions (CO2)_Agrifood Systems Waste Disposal	Emissions (N2O)_Agrifood Systems Waste Disposal	Emissions (CH4)_Burning - Crop residues	Emissions (N2O)_Burning - Crop residues	Emissions (N2O)_Crop Residues
0	Afghanistan	1991	28.7648	0.0000	0.3599	2.7606	0.0716	0.7906
1	Afghanistan	1992	29.9255	0.0000	0.3595	2.5314	0.0656	0.7416
2	Afghanistan	1993	31.3371	0.0000	0.3956	3.0714	0.0796	0.8710
3	Afghanistan	1994	32.7076	0.0000	0.4090	3.2408	0.0840	0.9134
4	Afghanistan	1995	33.8030	0.0000	0.4172	3.2016	0.0830	0.9201

(1573, 57)

	Domain	Area	Element	Item	Year	Unit	Value
0	Emissions intensities	Afghanistan	Production	Rice	1991	t	335000.0
1	Emissions intensities	Afghanistan	Production	Rice	1992	t	300000.0
2	Emissions intensities	Afghanistan	Production	Rice	1993	t	300000.0
3	Emissions intensities	Afghanistan	Production	Rice	1994	t	342000.0
4	Emissions intensities	Afghanistan	Production	Rice	1995	t	390000.0

(991, 7)

	Country	Year	Rice Production(Tonne)
0	Afghanistan	1991	335000.0
1	Afghanistan	1992	300000.0
2	Afghanistan	1993	300000.0
3	Afghanistan	1994	342000.0
4	Afghanistan	1995	390000.0

(991, 3)

	ObjectId	Country	ISO2	ISO3	Indicator	Unit	CTS Code	CTS Name	CTS Full Descriptor	1961	...	2014	2015	2016	2017
0	1	Afghanistan	AF	AFG	Temperature change with respect to a baseline ...	Degree Celsius	ECCS	Surface Temperature Change	Environment, Climate Change, Climate and Weath...	-0.126	...	0.521	1.204	1.612	1.642
1	2	Africa	NaN	AFRTMP	Temperature change with respect to a baseline ...	Degree Celsius	ECCS	Surface Temperature Change	Environment, Climate Change, Climate and Weath...	-0.017	...	1.013	1.190	1.392	1.180
2	3	Albania	AL	ALB	Temperature change with respect to a baseline ...	Degree Celsius	ECCS	Surface Temperature Change	Environment, Climate Change, Climate and Weath...	0.635	...	1.285	1.667	1.558	1.196

(236, 72)

	Country	Year	Temperature Index
0	Afghanistan	1991	-0.051
1	Africa	1991	0.375
2	Albania	1991	-0.280
3	Algeria	1991	0.102
4	American Samoa	1991	0.677

(7316, 3)

02c | DATA MERGING & PREPROCESSING

Inconsistent Country Name

eg. China, Hong Kong SAR , Hong Kong
Viet Nam, Vietnam
Iran (Islamic Republic of), Iran

Implement a fuzzy matching algorithm to standardize country names between the 3 datasets

```
all_countries = set(countries_pivot_df) | set(countries_merged_df) | set(countries_df_melted)

def map_country(country, country_list, threshold=90):
    match, score = process.extractOne(country, country_list)
    return match if score >= threshold else country

merged_df['Country'] = merged_df['Country'].apply(map_country, args=(all_countries,))
df_melted['Country'] = df_melted['Country'].apply(map_country, args=(all_countries,))
pivot_df['Country'] = pivot_df['Country'].apply(map_country, args=(all_countries,))
```

(Hong Kong), (China, Hong Kong SAR) -> Hong Kong

List of countries available in all datasets are different

Perform inner join on the "Country" and "Year" columns to include only the countries available in all three datasets.

```
result_df = pd.merge(merged_df, df_melted, pivot_df on=['Country', 'Year'], how='inner')
```

33 Asian Countries, Year: 1991-2021

Presencer of NaN values after merging the data

if (Year==2021)
Look at the values for the previous 5 years for the column.
if (prev_five_years.isna().all()) -> Impute 0
else -> Impute prev_five_years.mean(

if(Year!=2021)
Look at the values for the next 5 years for the column
if (next_five_years.isna().all()) -> Impute 0
else -> Impute next_five_years.mean(

RESULTS

	Country	Year	Emissions (CH4)_Agrifood Systems Waste Disposal	Emissions (CO2)_Agrifood Systems Waste Disposal	Emissions (N2O)_Agrifood Systems Waste Disposal	Emissions (CH4)_Burning - Crop residues	Emissions (N2O)_Burning - Crop residues	Emissions (N2O)_Crop Residues	Emissions (CO2)_Drained organic soils (CO2)	...	Rice Production(Tonne)	Temperature Index	Surface Temperature
0	Afghanistan	1991	28.7648	0.0000	0.3599	2.7606	0.0716	0.7906	0.0	...	335000.0	-0.051	decrease
1	Afghanistan	1992	29.9255	0.0000	0.3595	2.5314	0.0656	0.7416	0.0	...	300000.0	-0.212	decrease
2	Afghanistan	1993	31.3371	0.0000	0.3956	3.0714	0.0796	0.8710	0.0	...	300000.0	0.254	increase

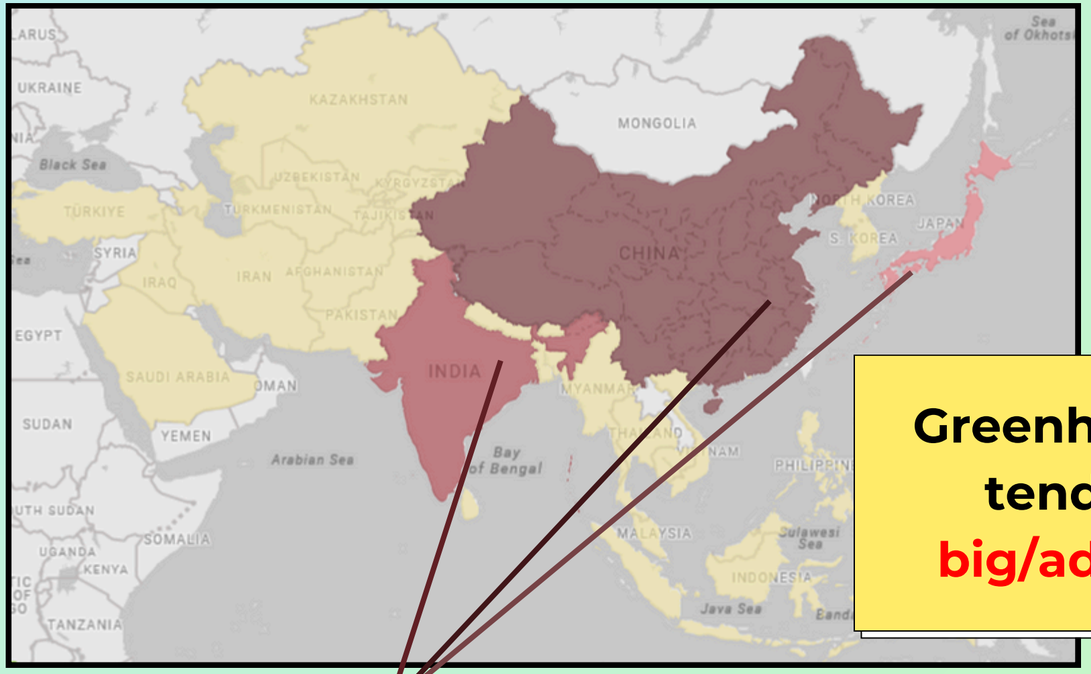
← Feature Enginering

Temperature index
"+ value" -> increase
"- value" -> decrease

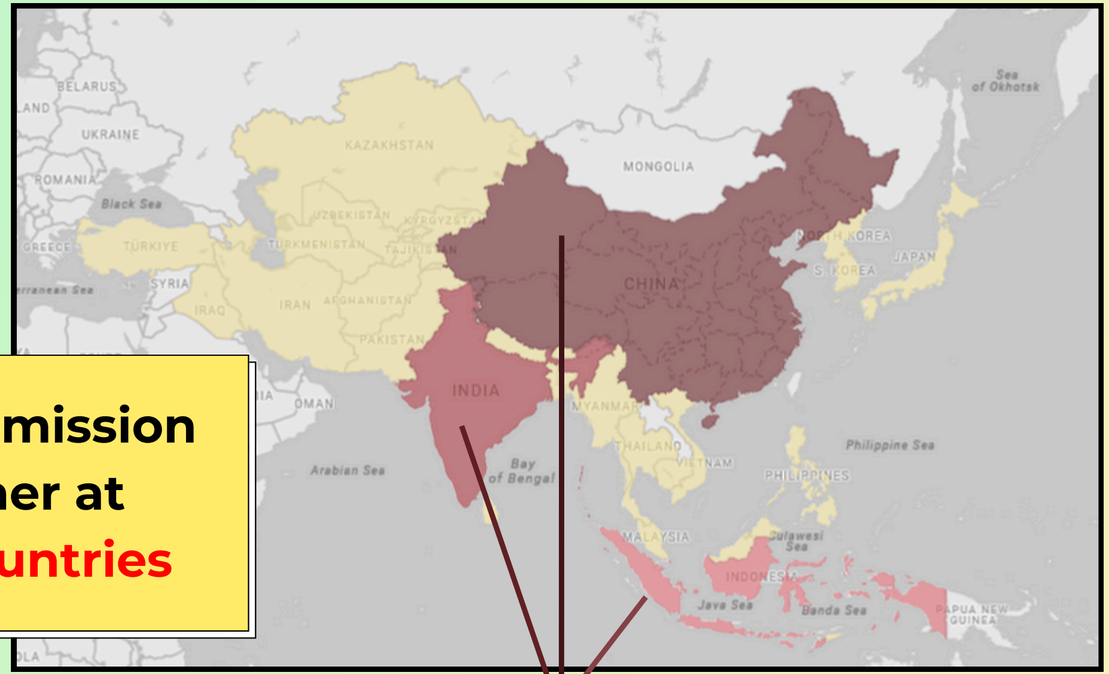
(985, 59)

03a | FACTORS IDENTIFICATION

TOP **Greenhouse Gas Emissions** Asian countries



TOP **Rice Production** Asian countries



1

Greenhouse gas emission tend to be higher at big/advanced countries

2

Greenhouse gas emission amount is highly affected by the agri-food production activities to increase rice production

TOP 1: CHINA
TOP 2: INDIA
TOP 3: JAPAN

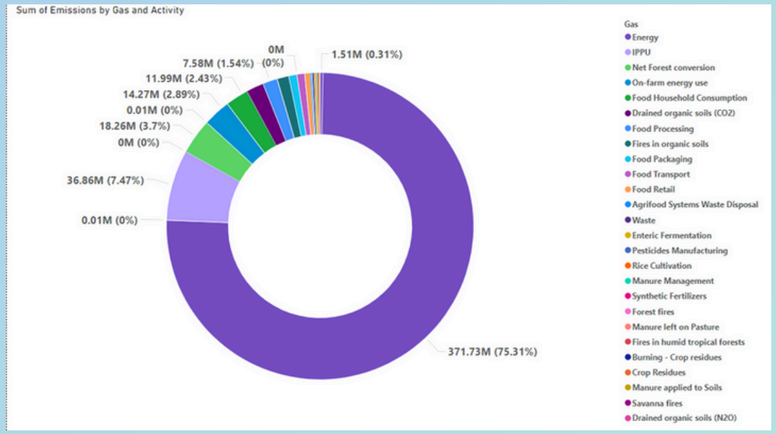
TOP 1: CHINA
TOP 2: INDIA
TOP 3: INDONESIA

03b FACTORS IDENTIFICATION

ENERGY GENERATION

emits 75% of the greenhouse gases

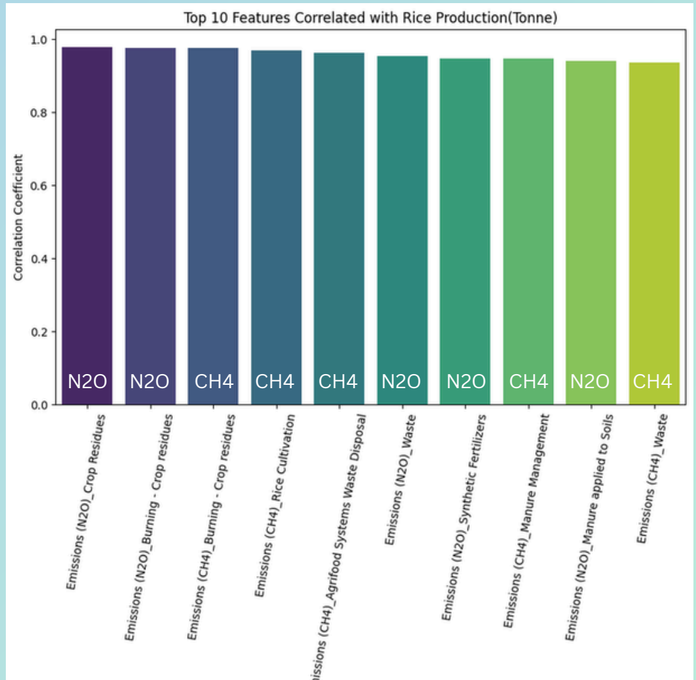
Energy generation refers to the burning of motor gasoline, gas-diesel oils, natural gas, liquefied petroleum gas, and to generate energy for energy-intensive processes for activities such as irrigation, fertilization and harvesting, .

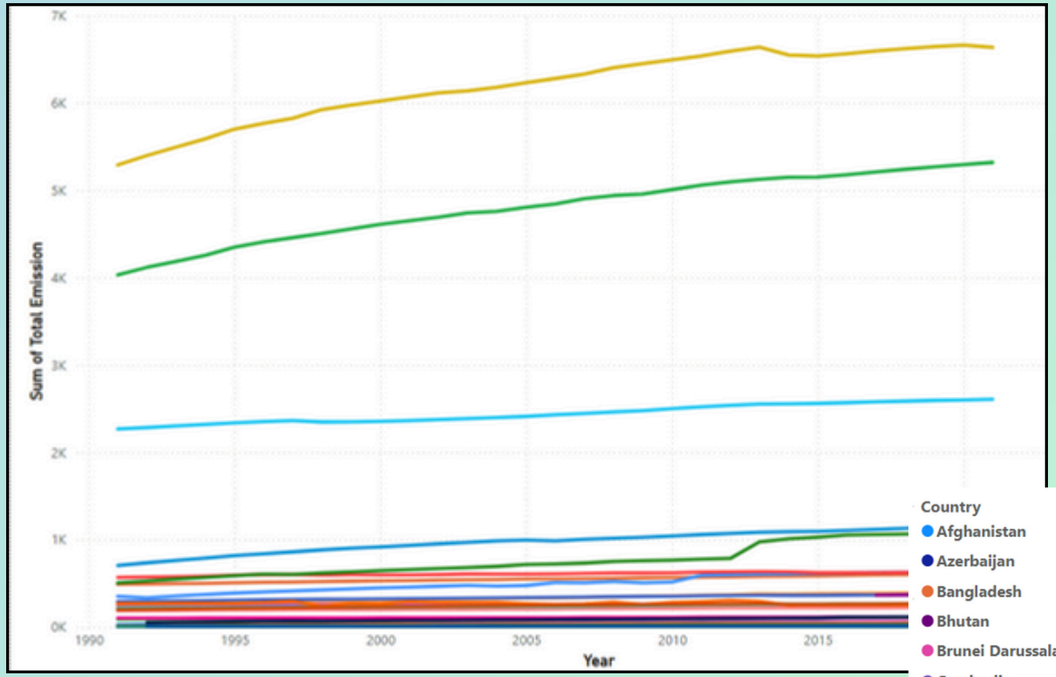


Most greenhouse gas generated by rice production are

N2O and CH4

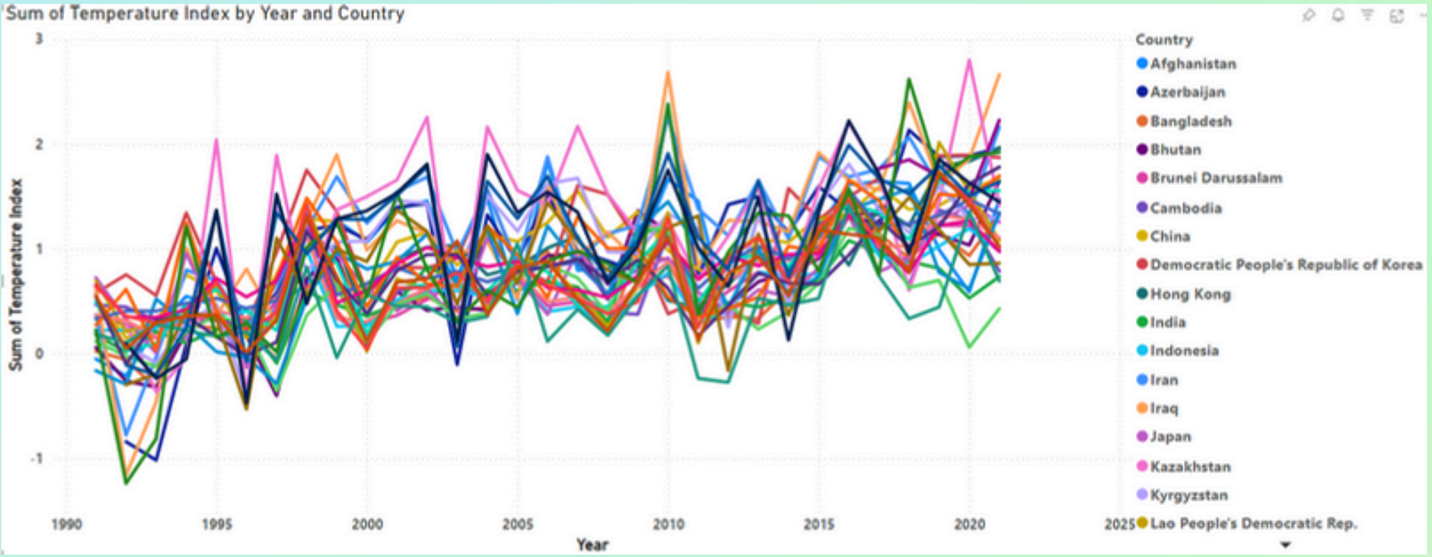
The releasing of (N2O) often due to burning crop residues ito quickly clear fields for the next planting season.





Greenhouse gases emission of each coutries are STEADILY INCREASING since 1991

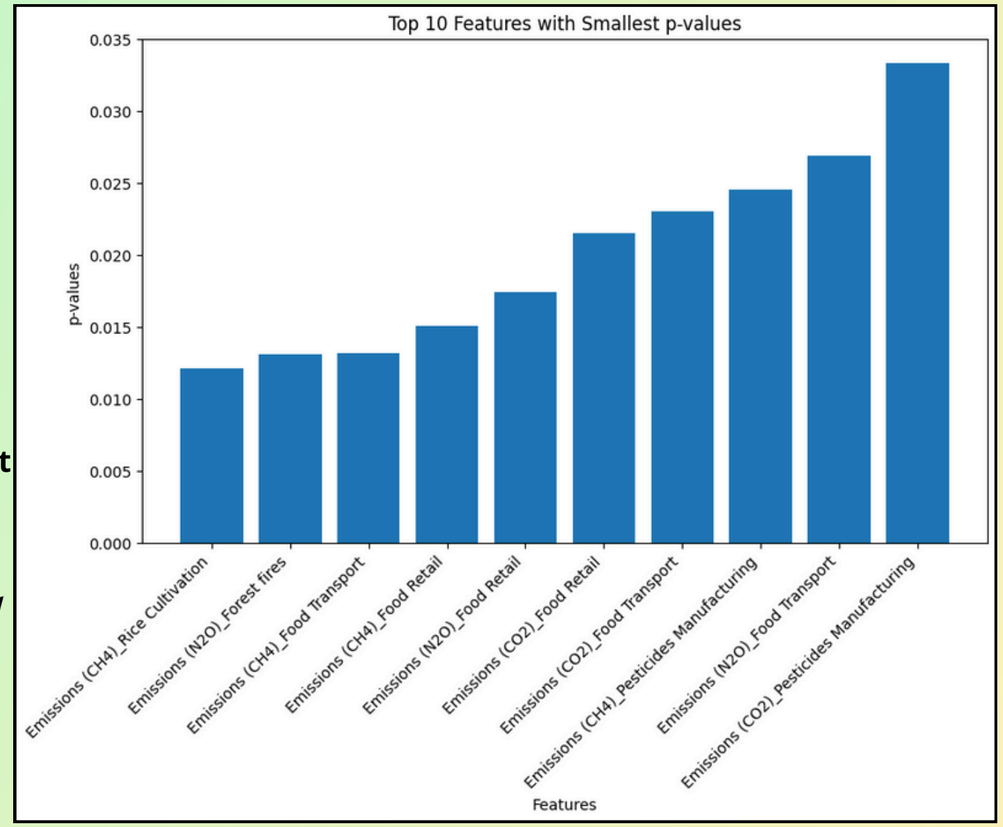
04a | EFFECT: TEMPERATURE CHANGE



Due to increasing emissions of greenhouse gases, the surface temperatures of all Asian countries are experiencing a **CONTINUOUSLY GROWING TREND**

Emissions of CH4 in Rice Cultivation is the key factor in affecting the **increasing / decreasing Surface Temperature**

Lower p-values suggest that the **feature may be associated with** variations in the **target variable** ('Surface Temperature')



One-way ANOVA Test

04b | CLASSIFICATION MODEL

To CLASSIFY THE EFFECT (TEMPERATURE CHANGE)

by emissions of greenhouse gases from agri-food activities

Random Forest Classifier

```
# Random Forest Classifier
print('Random Forest Classifier')
rf_model = RandomForestClassifier()
rf_model.fit(oversampled_X_train, oversampled_y_train)
rf_predictions = rf_model.predict(X_test)
```

Accuracy: 0.957 Recall: 0.954 Log-loss: 0.107
Precision: 0.954 F1 score: 0.954 ROC-AUC: 0.994

Neural Network Classifier

```
# Neural Network Classifier
print('Neural Network Classifier')
nn_model = MLPClassifier()
nn_model.fit(oversampled_X_train, oversampled_y_train)
nn_predictions = nn_model.predict(X_test)
```

Accuracy: 0.765 Recall: 0.690 Log-loss: 4.689
Precision: 0.779 F1 score: 0.731 ROC-AUC: 0.823

Decision Tree Classfier

```
# Decision Tree Classifier
print("Decision Tree Classifier")
dt_classifier = DecisionTreeClassifier(random_state=42)
dt_classifier.fit(oversampled_X_train, oversampled_y_train)
dt_predictions = dt_classifier.predict(X_test)
```

Accuracy: 0.955 Recall: 0.954 Log-loss: 1.634
Precision: 0.949 F1 score: 0.951 ROC-AUC: 0.955

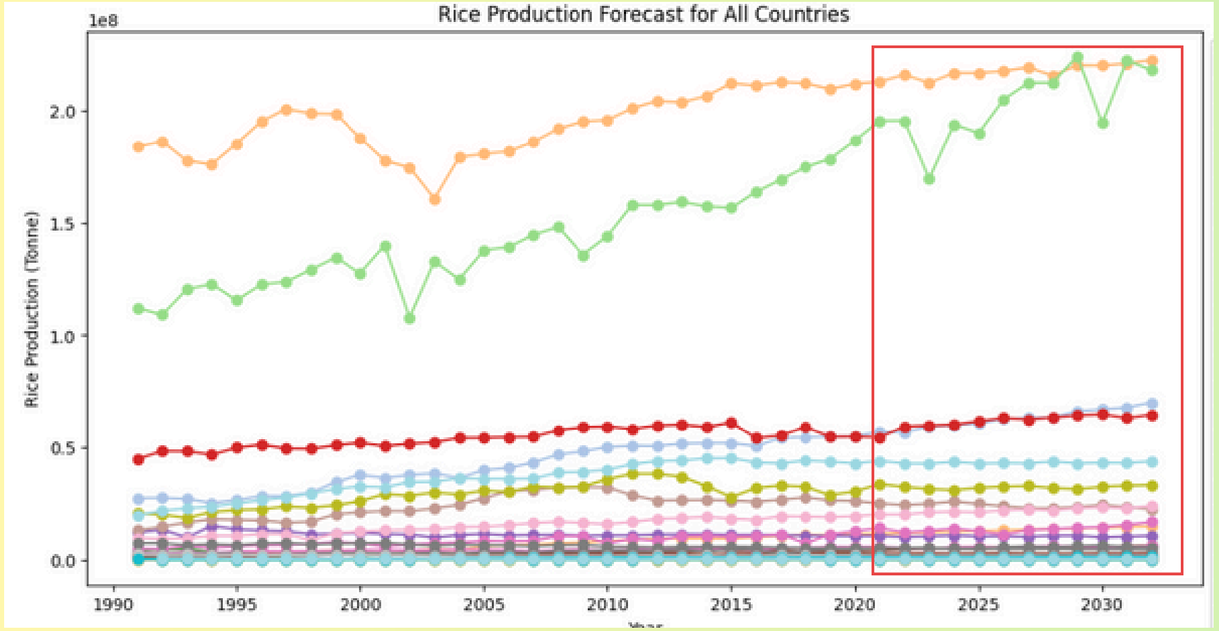
K-Nearest Neighbours Classifier

```
# KNN classifier
print('K-Nearest Neighbours Classifier')
knn_classifier = KNeighborsClassifier()
knn_classifier.fit(oversampled_X_train, oversampled_y_train)
knn_predictions = knn_classifier.predict(X_test)
```

Accuracy: 0.875 Recall: 0.799 Log-loss: 1.239
Precision: 0.921 F1 score: 0.855 ROC-AUC: 0.936

05a | TIME SERIES FORECASTING

Exponential Smoothing

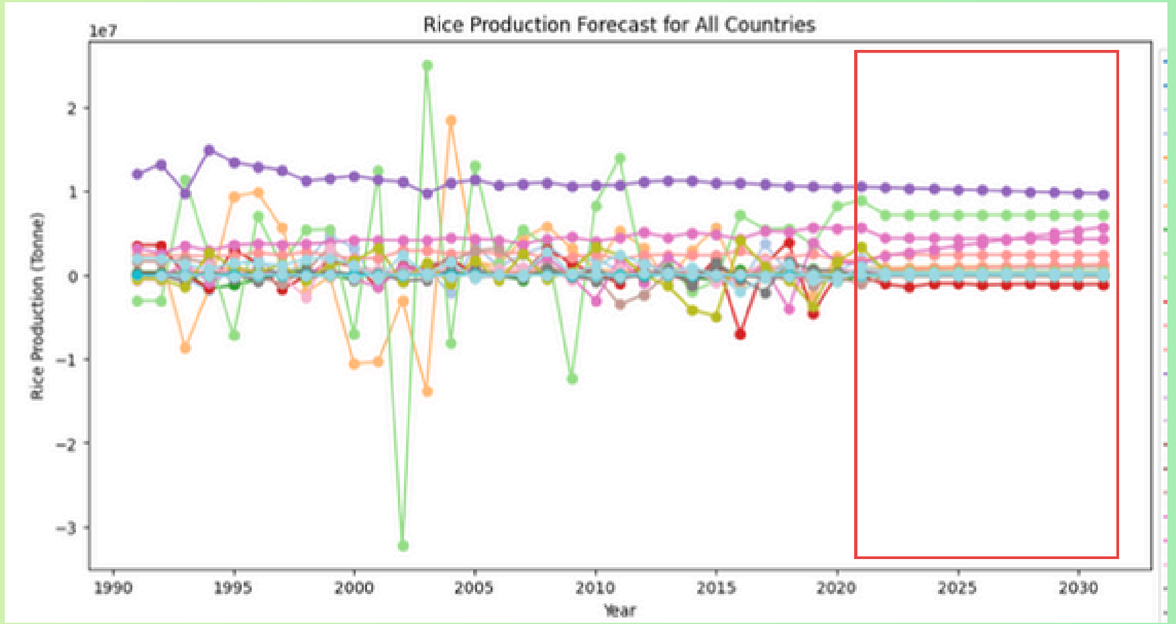


Country	Trend	Seasonal	Seasonal Period	RMSE
Afghanistan	add	add	8	3.187100e+09
Azerbaijan	add	add	9	1.315274e+07
Bangladesh	mul	add	6	7.722352e+11
Bhutan	add	mul	9	3.431920e+08
Brunei Darussalam	mul	mul	9	3.982861e+04
Cambodia	add	add	3	4.017503e+10
China	mul	add	5	5.882385e+12
Democratic People's Republic of Korea	add	mul	3	6.056946e+10
Hong Kong	add	add	9	1.489645e-05
India	mul	mul	7	3.810285e+13
Indonesia	mul	add	10	1.203843e+13
Iran	add	add	10	1.282234e+11
Iraq	mul	mul	7	2.875175e+10
Japan	mul	mul	8	6.056669e+10
Kazakhstan	mul	mul	8	1.728287e+09
Kyrgyzstan	add	mul	2	1.806660e+06
Lao People's Democratic Rep.	add	mul	8	4.284476e+11
Malaysia	add	add	10	3.203904e+10
Myanmar	mul	mul	6	2.496757e+11
Nepal	mul	mul	8	1.304988e+11
Pakistan	mul	mul	8	2.345374e+12
Philippines	add	add	3	1.681540e+11
Republic of Korea	mul	add	3	5.541369e+10
Sri Lanka	mul	mul	7	6.017216e+11
Taiwan Province of China	mul	add	9	2.206863e+10
Tajikistan	add	add	2	3.526251e+08
Thailand	mul	add	5	2.084163e+12
Timor-Leste	add	add	7	6.504350e+07
Turkiye	add	add	3	5.552162e+08
Turkmenistan	mul	mul	2	2.398085e+08
Uzbekistan	add	add	9	8.341672e+09
Vietnam	mul	add	4	1.034385e+12

Rice production
INCREASE STEADILY
in the future 10 years

Higher RSME:
More deviation from the actual value

Autoregressive Integrated Moving Average (ARIMA)



	Country	p	d	q	RMSE
0	Afghanistan	-1	0	0	3.174692e+04
1	Azerbaijan	1	0	1	4.396287e+03
2	Bangladesh	-1	0	-1	1.441319e+06
3	Bhutan	-1	2	-1	1.437440e+04
4	Brunei Darussalam	0	0	0	4.403683e+02
5	Cambodia	-1	2	1	1.957320e+05
6	China	0	0	1	1.771009e+06
7	Democratic People's Republic of Korea	0	1	1	4.326212e+05
8	Hong Kong	0	2	0	4.470157e-03
9	India	-1	1	-1	2.116858e+06
10	Indonesia	1	1	0	1.926321e+06
11	Iran	-1	0	-1	5.255706e+05
12	Iraq	-1	2	-1	1.674837e+05
13	Japan	1	2	-1	1.285671e+05
14	Kazakhstan	0	2	-1	3.988833e+04
15	Kyrgyzstan	0	0	1	1.380251e+03
16	Lao People's Democratic Rep.	-1	2	0	2.763486e+05
17	Malaysia	-1	1	1	1.443942e+05
18	Myanmar	-1	1	1	1.000989e+06
19	Nepal	1	0	0	9.890979e+05
20	Pakistan	-1	2	-1	2.488154e+06
21	Philippines	-1	0	0	5.018578e+05
22	Republic of Korea	-1	0	0	2.948063e+05
23	Sri Lanka	1	0	-1	1.080947e+06
24	Taiwan Province of China	1	0	1	1.495311e+05
25	Tajikistan	0	0	1	1.345951e+04
26	Thailand	1	0	-1	2.341221e+06
27	Timor-Leste	1	0	1	3.806306e+03
28	Turkiye	-1	1	1	3.342591e+04
29	Turkmenistan	1	2	1	1.925753e+04
30	Uzbekistan	-1	0	1	8.624972e+04
31	Vietnam	-1	1	1	9.032890e+05

Rice production
REMAIN CONSTANT
in the future 10 years

Lower RSME:
Lower deviation from the actual value

05b | TIME SERIES FORECASTING

HOWEVER...

Exponential Smoothing

LEVEL
Weight given to the most recent observation in calculating the smoothed value.

TREND
Weight given to the trend component in calculating the smoothed value.

SEASONALITY
Weight given to the seasonal component in calculating the smoothed value.

- **More robust to noise** especially if data is not large
- Does not require differencing
- Produce smoother forecasts that **capture trends and more naturally**

ARIMA

AUTOREGRESSIVE
Number of lagged observations included in the model for autoregressive component.

INTEGRATED
Degree of differencing applied to the time series to achieve stationarity.

MOVING AVERAGE
Number of lagged forecast errors included in the model for moving average component.

- **Less robust to noise** if data is not large
- Require differencing which introduce **negative value** to the strictly positive data
- Tend to produce straight lines because they predict future values **based on a linear combination of past observations.**

After understanding the
UNEVITABLE RISE IN RICE PRODUCTION ,
what can we do?

06 RECOMMENDATIONS

Renewable Energy Integration

- Solar power, wind power, biogas plants
- Reducing greenhouse gas emissions during burning of fossil fuels

Precision Agriculture

- Implement AI and machine learning for crop management
- Optimize production and minimize energy consumptions and gas emission

Greenhouse Gas Capture Technologies

- Carbon sequestration and methane digesters
- Storing greenhouse gas emission for recycle to mitigate climate change

Cold Chain Logistics Improvement

- Reduce food spoilage and energy consumption

07 CONCLUSION

SIGNIFICANT FINDINGS

Big or advanced countries such as China, India and Japan has the largest greenhouse gas emission amount

It is proved that agri-food production amounts has significant impact on greenhouse gas emission

Renewable energy implementation need prompt attention as energy consumption contributes most to greenhouse gas

Emission of CH4 in Rice Cultivation is one of the major features that lead to rise in temperature

Agri-food production of each country is expected to be steadily increasing in the future 10 years

Recommendations have been proposed to mitigate greenhouse gas emission due to the increasing agri-food production

LIMITATIONS

Not all Asian countries are involve in this research due to lack of dataset

Imbalanced class: ‘decrease’ class in the Surface Temperature column is underrepresented compared to the ‘increase’ class, in which we mitigate SMOTE to balance the classes

Rice production time series forecasting assume without huge climate change or disasters