

Name: **Hafsa Jehanzaib**

Class: **BSCS**

Section: **6B**

Roll No: **18122035**

OBJECTIVE:

Prediction of Economic growth Forecast GDP(Gross Domestic Product) trend using GDP(current US\$),Inflation ,GDP deflator.

Introduction:

The dataset that is being examined offers a wide range of socio-economic indicators that are essential for comprehending and assessing the state of development in a certain area or nation. This dataset, which has columns covering demographic, economic, and developmental variables, provides a wealth of information to explore the trends, patterns, and connections that influence a country's course throughout time.

Background:

The basis for comprehending a country's economic performance and stability is laid by economic statistics such as GDP, total reserves, gross national expenditure, and inflation rates. Particularly, GDP serves as a gauge of the state of the economy, while national spending and inflation rates provide insight into fiscal policies and methods for stimulating the economy.

Data Sources:

The dataset compiles data from a number of reliable sources, such as kaggle (the World Bank) and other statistical, economic data.

Data Preprocessing

✓ Handling Missing Values:

Identify and address missing values across all columns. Various techniques, such as or removal of rows/columns with significant missing data, were based on the nature and extent of missingness in each attribute. This ensured dataset's integrity while retaining valuable information.

- ✓ Binning
- ✓ Remove Duplicates
- ✓ Sorting

Modelling

Various analytical methods and algorithms were employed to extract meaningful insights from the dataset. These methods include but are not limited to:

- ✓ **SVM model**
- ✓ **Linear regression:** Utilized to understand relationships between dependent and independent variables, such as predicting GDP based on economic indicators.

Time series forecasting: Employed to analyze trends, seasonality, and patterns in the dataset, especially when dealing with temporal data like population growth over years.

Evaluation Metrics:

Performance metrics utilized to evaluate the models' effectiveness. Metrics such as Mean Squared Error (MSE) were considered based on the nature of the modeling task (e.g., regression random forest).

RESULTS:

Local Repository/final project* - RapidMiner Studio Educational 10.2.000 @ Ele

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Interactive Analysis

Find data, operators... etc All Studio

Repository

- Training Resources (connected)
- Samples
- Community Samples (connected)
- Local Repository (Local)
- DB (Legacy)

Operators

retrie

- Data Access (1)
 - Retrieve
- Blending (1)
 - Attributes (1)
 - Types (1)
 - Guess Types

No results were found.

Process

Process

90%

Read CSV → Remove Duplicates → Replace Missing Value... → Sort → Discretize

Retrieve → Split Data → Set Role → Replace Missing Value... → Linear Regression → Apply Model

Select Attributes

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

Activate Wisdom of Crowds

Parameters

Process

logverbosity: init

logfile:

Show advanced parameters

Change compatibility (10.2.000)

Help

Process

RapidMiner Studio Core

Synopsis

The root operator which is the outer most operator of every process.

Local Repository/final project* - RapidMiner Studio Educational 10.2.000 @ Ele

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Interactive Analysis

Find data, operators... etc All Studio

Result History

LinearRegression (Linear Regression) ExampleSet (Apply Model)

LinearRegression

Data

1993.421

Description

Annotations

<new process*> - RapidMiner Studio Educational 10.2.000 @ Ele

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Interactive Analysis

Find data, operators... etc All Studio

Repository

Import Data

- Training Resources (connected)
- Samples
- Community Samples (connected)
- Local Repository (Local)
- DB (Legacy)

Operators

missing

- Replace All Missing
- Handle Unknown Values
- Quality Measures
- Modeling (1)
 - Time Series (1)
 - Transformation (1)

We found "Database Envy" in the Marketplace. [Show me!](#)

Process

Process

75%

Read CSV → Split Data → Replace Missing Values → SVM → Apply Model → Performance

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

Activate Wisdom of Crowds

Parameters

Process

logverbosity: init

logfile:

Show advanced parameters

Change compatibility (10.2.000)

Help

Process

RapidMiner Studio Core

Synopsis

The root operator which is the outer most operator of every process.

<new process*> - RapidMiner Studio Educational 10.2.000 @ Ele

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Interactive Analysis

Find data, operators... etc All Studio

Result History

Kernel Model (SVM) ExampleSet (Apply Model) PerformanceVector (Performance)

Kernel Model

Description

Total number of Support Vectors: 44
Bias (offset): 86538763600.341

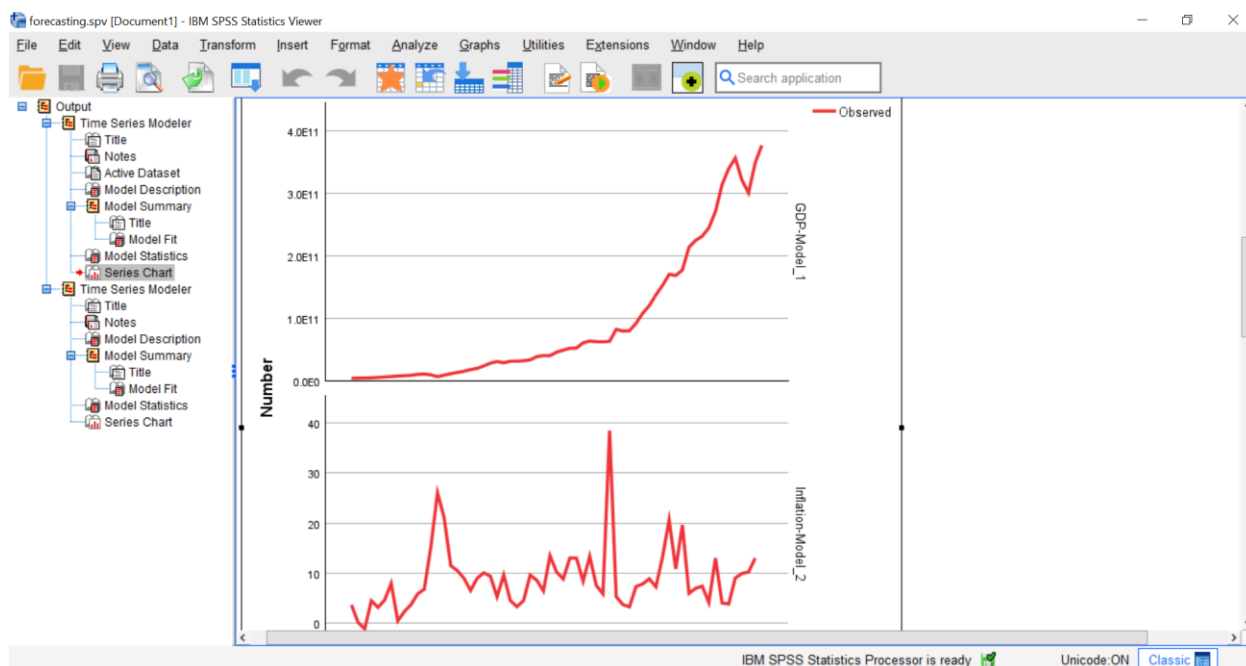
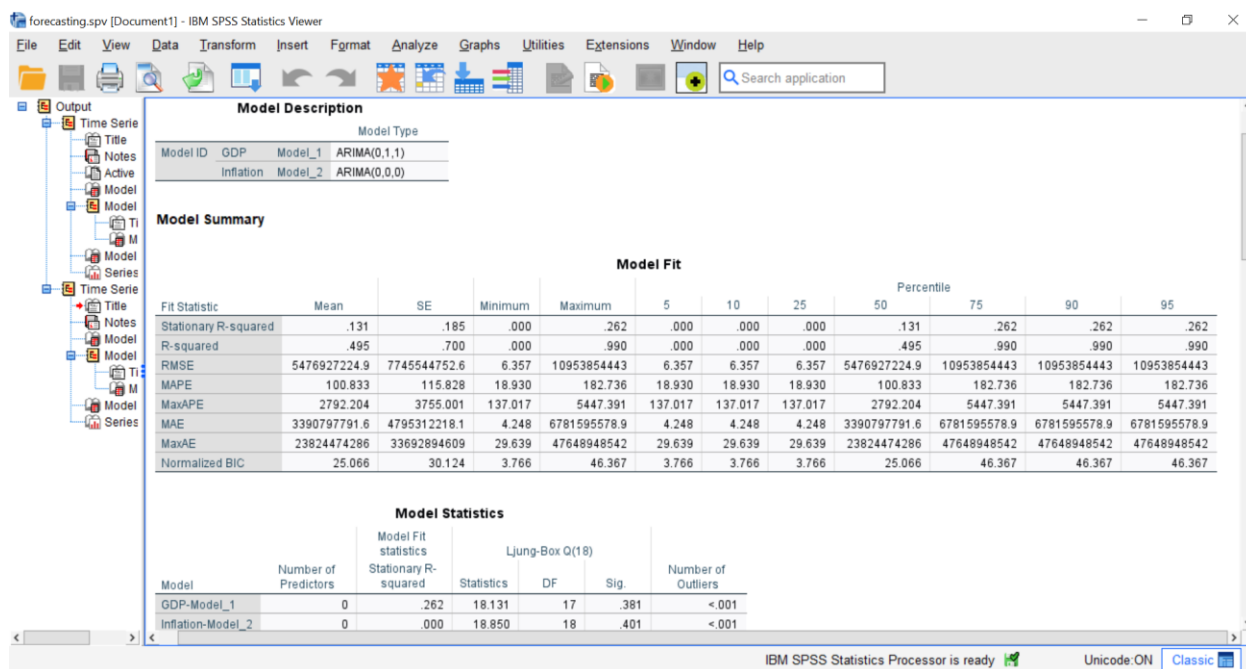
Weight Table

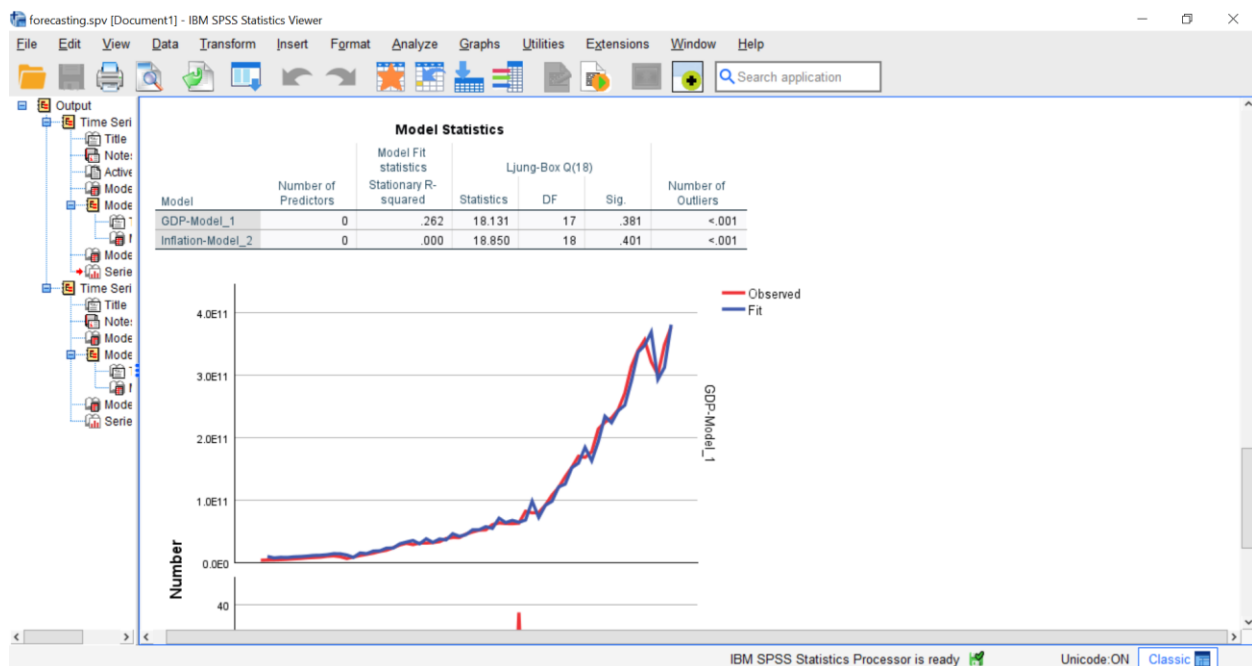
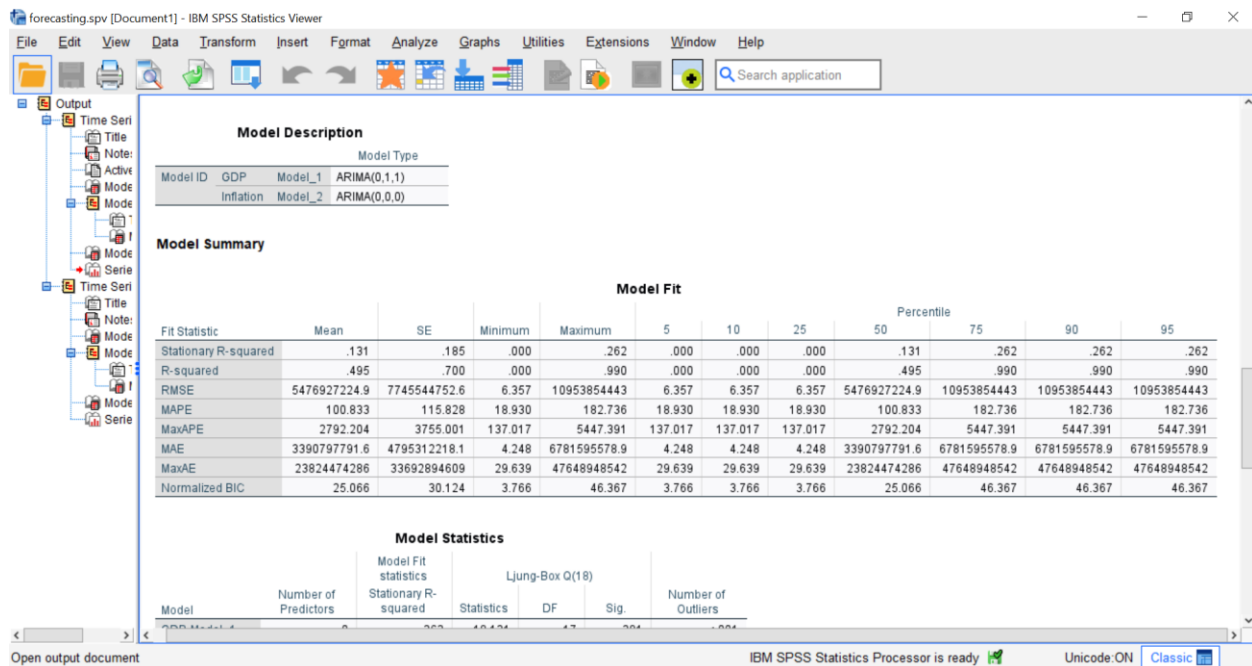
Weight Visualization

Support Vector Table

Support Vector Visualization

w[Year] = 1.962
w[Urban population] = 1.941
w[Rural population] = 1.999
w[Population, female] = 1.970
w[Population, male (% of total population)] = -1.846
w[Population, total] = 1.978
w[Population growth (annual %)] = -1.234
w[Birth rate, crude (per 1,000 people)] = -1.887
w[Mortality rate, infant (per 1,000 live births)] = -1.926
w[Inflation, GDP deflator (annual %)] = 0.354
w[GNI (current US\$)] = 1.514
w[Gross national expenditure (current US\$)] = 1.495
w[Total reserves (includes gold, current US\$)] = 1.383
w[Services, value added (current US\$)] = 1.495
w[Merchandise exports (current US\$)] = 1.796
w[Merchandise imports (current US\$)] = 1.504
w[Military expenditure (current US\$)] = 1.649
w[Adjusted savings: education expenditure (current US\$)] = 1.388
w[Food production index (2014-2016 = 100)] = 1.803





*final timeseies forecasting.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Visible: 8 of 8 Variables

	GDP	Inflation	YEAR	DATE	Predict ed_GD P_Mod.	Predict ed_Infl ation_...	Predict ed_GD P_Mod.	Predict ed_Infl ation_...	var	var	var	var	var	var
1				1960	1960									
2	3749265015			1961	1961			-7.E+10						
3	4118647627		4	1962	1962	1.E+10	9	-6.E+10	7					
4	4310163797		0	1963	1963	7.E+9	9	-6.E+10	7					
5	4630827383		-1	1964	1964	8.E+9	9	-5.E+10	7					
6	5204955901		4	1965	1965	8.E+9	9	-5.E+10	7					
7	5929231415		3	1966	1966	9.E+9	9	-4.E+10	7					
8	6561108778		5	1967	1967	1.E+10	9	-4.E+10	7					
9	7464510710		8	1968	1968	1.E+10	9	-3.E+10	7					
10	8041999160		0	1969	1969	1.E+10	9	-3.E+10	7					
11	8683116338		2	1970	1970	1.E+10	9	-2.E+10	7					
12	10027509450		4	1971	1971	1.E+10	9	-2.E+10	7					
13	10665896682		6	1972	1972	1.E+10	9	-1.E+10	7					
14	9415016360		7	1973	1973	1.E+10	9	-5.E+9	7					
15	6383429490		15	1974	1974	1.E+10	9	-40101122	8					
16	8899191919		26	1975	1975	8.E+9	9	5.E+9	8					
17	11230606061		21	1976	1976	2.E+10	9	1.E+10	8					
18	13168080808		12	1977	1977	1.E+10	9	2.E+10	8					
19	15126060606		11	1978	1978	2.E+10	9	2.E+10	8					
20	17811515152		9	1979	1979	2.E+10	9	3.E+10	8					
21	19688383838		7	1980	1980	2.E+10	9	3.E+10	8					

*final timeseies forecasting.sav [DataSet1] - IBM SPSS Statistics Data Editor





File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Visible: 8 of 8 Variables

	GDP	Inflation	YEAR	DATE	Predict ed_GD P_Mod.	Predict ed_Infl ation_...	Predict ed_GD P_Mod.	Predict ed_Infl ation_...	var	var	var	var	var	var
22	23654444444		9	1981	1981	2.E+10	9	4.E+10	8					
23	28100606061		10	1982	1982	3.E+10	9	4.E+10	8					
24	30725971564		9	1983	1983	3.E+10	9	5.E+10	8					
25	29691889764		5	1984	1984	4.E+10	9	5.E+10	8					
26	31151825467		10	1985	1985	3.E+10	9	6.E+10	8					
27	31144920844		5	1986	1986	4.E+10	9	6.E+10	8					
28	31899070056		3	1987	1987	3.E+10	9	7.E+10	8					
29	33351529275		5	1988	1988	4.E+10	9	7.E+10	9					
30	38472742808		10	1989	1989	4.E+10	9	8.E+10	9					
31	401711105580		9	1990	1990	5.E+10	9	8.E+10	9					
32	40010423970		6	1991	1991	4.E+10	9	9.E+10	9					
33	45625336680		13	1992	1992	4.E+10	9	9.E+10	9					
34	48884672605		10	1993	1993	5.E+10	9	1.E+11	9					
35	51809999353		9	1994	1994	5.E+10	9	1.E+11	9					
36	52293471393		13	1995	1995	6.E+10	9	1.E+11	9					
37	60636071684		13	1996	1996	5.E+10	9	1.E+11	9					
38	63320170084		8	1997	1997	7.E+10	9	1.E+11	9					
39	62433340468		13	1998	1998	6.E+10	9	1.E+11	9					
40	62191955814		8	1999	1999	7.E+10	9	1.E+11	9					
41	62973856844		6	2000	2000	6.E+10	9	1.E+11	9					
42	82017743416		39	2001	2001	7.E+10	9	1.E+11	10					

Overview Data View Variable View

IBM SPSS Statistics Processor is ready Unicode:ON Classic

	 GDP	 Inflation	 YEAR	 DATE	Predict ed_GD P_Mod.	Predict ed_infl ation_	Predict ed_GD P_Mod.	Predict ed_infl ation_	var	var	var	var	var	var
44	79904985385		4	2003	2003	7 E+10	9	2 E+11	10					
45	91760542940		3	2004	2004	9 E+10	9	2 E+11	10					
46	107760000000		7	2005	2005	1 E+11	9	2 E+11	10					
47	120055000000		8	2006	2006	1 E+11	9	2 E+11	10					
48	137264000000		9	2007	2007	1 E+11	9	2 E+11	10					
49	152386000000		7	2008	2008	2 E+11	9	2 E+11	10					
50	170078000000		13	2009	2009	2 E+11	9	2 E+11	10					
51	168153000000		21	2010	2010	2 E+11	9	2 E+11	10					
52	177166000000		11	2011	2011	2 E+11	9	2 E+11	10					
53	213588000000		20	2012	2012	2 E+11	9	2 E+11	10					
54	224384000000		6	2013	2013	2 E+11	9	2 E+11	10					
55	231218000000		7	2014	2014	2 E+11	9	2 E+11	10					
56	244361000000		7	2015	2015	2 E+11	9	2 E+11	11					
57	270556000000		4	2016	2016	3 E+11	9	2 E+11	11					
58	313630000000		13	2017	2017	3 E+11	9	2 E+11	11					
59	339206000000		4	2018	2018	3 E+11	9	2 E+11	11					
60	356128000000		4	2019	2019	3 E+11	9	2 E+11	11					
61	320909000000		9	2020	2020	4 E+11	9	2 E+11	11					
62	300426000000		10	2021	2021	3 E+11	9	2 E+11	11					
63	348263000000		10	2022	2022	3 E+11	9	3 E+11	11					
64	376533000000		13	2023	2023	4 E+11	9	3 E+11	11					