The Impact on Worldwide of Labour Force Employment Since Last 19 Years (2000-2019) by using Different Classification Techniques of Pattern Recognition

Gajanand A. Boywar[1], Chandrakant Patil[1], Pratik S. Jaiswal[1], Mukta G. Dhopeshwarkar[1]

[1]Department of Computer Science & I.T, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra - 431004 India.

---***---

Abstract. The acknowledgment of work power interest rate and joblessness rate has turned into a significant part pondering in Employments Sector. Example acknowledgment turns into a significant part as the entire investigation comprise of statistics. In the present examination, distinctive classifier of example recognition have been utilized. In which we have found extraordinary result by utilizing DecisionTree classifier, HoeffdingTree classifier, J48 classifier, LMT classifier, RandomForest classifier, RandomTree classifier, and in conclusion REPTree classifier i.e., 33.4298%, 39.725%, 82.1274%, 79.3054%, 98.4986%, 98.4624%, 66.4797% individually. When we apply DecisionTree classifier the result is poor interestingly with other classifier for example 33.4298%. what's more, Other hand the RandomForest classifier result is high conversely with other classifier for example 98.4986%. in case on the off chance that we apply classifier of same kind data later on, we could extraordinary results by the utilization of above classifier. Just DecisionTree classifier will be considered as uncommon. **Keyword.** Tree Classifier, DecisionTree, HoeffdingTree, J48, LMT, RandomForest, RandomTree, and REPTree.

1   Introduction

International Labor Organization (ILO) programmed on labor force participation rate and unemployment rate is part of large international effort on demographic estimated and projection to which several UN agencies contribute. The principle target of ILO program is to give individual from States, International organizations and the general population everywhere with the most extensive, itemized and practically identical work and Unemployment of the work power for nations and regions, the world all in all and its primary topographical area. the essential information are single-year work power interest rates by sex (percent)group and populace matured 15 years and over[9].

2   Data Information

The work control adventure dataset gather the data over the word by assembled in region cunning, nation wise and Area skillful. The dataset in collected between years 2000 to 2019[8-12]. In addition, the dataset pace of energy of utilized and jobless by sex (percent). The given break down dataset is amass from International Labor Organization(ILO),Geneva, Key Indicators of the Labor Market(KILM ninth discharge) and the ILOSTAT database, last got to January 2019. Additionally, Population developed 15 years and over, with the exception of if for the most part footnoted.

1. Information avoids Armenia, Azerbaijan, Cyprus, Georgia, Israel and Turkey.

2. Caucasus alludes to Armenia, Azerbaijan, Cyprus, Georgia, Israel furthermore, Turkey.

3. Populace matured 16 years and over.

4. Populace matured 14 years and over.

5. Occupant populace (by right).

6. Barring the institutional populace.

7. For factual purposes, the information for China do exclude those for the Hong Kong Special Administrative Region (Hong Kong SAR), Macao Special Administrative Region (Macao SAR) and TaiwanProvince of China.

8. Populace matured 15 to 69 years.

9. Populace matured 16 to 65 years.

10. Populace matured 15 to 74 years.

11. Nationals, occupants.

12. Populace matured 18 to 64 years.

13. Populace matured 15 to 64 years.

14. People present (true).

15. True populace.

16. Populace matured 17 years and over.

17. Fundamental city or metropolitan zone.

18. Excluding a few territories.

3   Methodology

Right off the bat, the way toward acquiring and fusing model information has been significantly changed so as to completely incorporated the dataset assemble from United Nations static division. At that point it was not in line way, so we need that to clean the database base in an effective way. For that first, we are perfect the no of occasions semantic goofs which are shown in the database classification[13]. Beginning their ahead, we refreshed all characteristics in a certifiable way. When we have done the whole above database cleaning process the database is set up to perform specific activities on it.
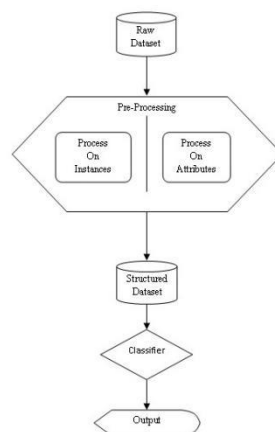

Fig.1. Illustration of Dataset process

DecisionStump is usually used in conjunction with a boosting algorithm. Does regreesion(based on mean-squared error) or classification(based on entopy). Missing is treated as separate value[1]. A HoeffdingTree (VFDT) is an incremental, anytime decision tree induction algorithm that is capable of learning from massive data streams, assuming that the distribution generating examples does not change over time[2]. J48 classes for generating a pruned or unpruned C4 which based on 5 decision tree[3]. LMT is used to classifier for building 'logostic model trees'. Which are classification trees with logistic regression functions at the leaves[4]. The RandomForest classifier used for create class for constructing a forest of random trees[5]. RandomTree used to create class for constructing a tree that considers chosen attributes at each node. Whoch performs no pruning. Also has an option to allow estimation of class probabilities(or target mean in the regression case) based on hold-out set (backfitting)[6]. The REPTree is a fast decision tree learner which builds a decision/regression tree using information gain/variance and prune it using reduced-error pruning[7].

4   Implementation

In the Labour Force Participation rate and unemployment rate(LFPRAUR) dataset the 5528 number of occasions subsequent to apply this classifier, the case in grouped in a different result. In DecisionStump classifier 33.4298% , a HoeffdingTree classifier is 39.725%, the J48 classifier is 62.1274%, the LMT classifier is 79.3054, the RandomForest is 98.4986%, the RandomTree classifier is 98.4624%,and at the last is REPTree Classifier is 66.4797%.

Table1. Result of DecisionStump

	No. Of instance	Result
Correctly Classified Instances	1848	33.4298%
Incorrectly Classified Instances	3680	66.5702%
Kappa statistic	0.2001	
Mean absolute error	0.2273	
Root mean squared error	0.3371	
Relative absolute error	81.8315%	
Root relative squared error	90.4608%	
Total Number of Instances	5528	

After performing the DecisionStump classifier operation on dataset with 5528 number the result is shows as in table1 which state that the Correctly Classified Instances 1848 which is in 33.4298 %, Incorrectly Classified   Instances 3680 which is   66.5702 %, Kappa statistic is 0.2001, Mean absolute error is 0.2273,Root mean squared error0.3371, Relative absolute error is 81.8315 %, Root relative squared error is 90.4608 %.

Table1.1. Confusion Matrix of DecisionStump

a	b	c	d	e	f
922	1	0	0	0	0
5	926	0	0	0	0
916	0	0	0	0	0
5	916	0	0	0	0
803	113	0	0	0	0
17	904	0	0	0	0

Where, a = Labour force participation Total, b = Unemployment rate Total, c = Labour force participation Male, d = Unemployment rate Male, e = Labour force participation   Female, f = Unemployment rate Female.

Table2. Result of HoeffdingTree

	No. Of instance	Result
Correctly Classified Instances	2196	39.725%
Incorrectly Classified Instances	3332	60.275%
Kappa statistic	0.2759	
Mean absolute error	0.234	
Root mean squared error	0.3374	
Relative absolute error	84.2287%	
Root relative squared error	90.5386%	
Total Number of Instances	5528	

Subsequent to playing out the HoeffdingTree classifier activity on dataset with 5528 number the outcome is appears as in table2 which express that the Correctly Classified Instances 2196 which is in 39.725 %, Incorrectly Classified Instances 3332 which is 60.275 %, Kappa measurement is 0.2759, Mean outright mistake is 0.234, Root mean squared blunder 0.3374, Relative supreme blunder is 84.2287 %, Root relative squared blunder is 90.5386 %.

Table2.1. Confusion Matrix of HoeffdingTree

a	b	c	d	e	f
719	5	62	0	137	0
1	927	0	0	3	0
497	1	376	0	42	0
1	918	0	0	2	0
601	119	22	0	174	0
9	906	0	0	6	0

Where, a = Labour force participation Total, b = Unemployment rate Total, c = Labour force participation Male, d = Unemployment rate Male, e = Labour force participation   Female, f = Unemployment rate Female.

Table3. Result of J48

	No. Of instance	Result
Correctly Classified Instances	4540	82.1274%
Incorrectly Classified Instances	988	17.8726%
Kappa statistic	0.7855	
Mean absolute error	0.07	
Root mean squared error	0.187	
Relative absolute error	25.1833%	
Root relative squared error	50.183%	
Total Number of Instances	5528	

In the wake of playing out the J48 classifier activity on dataset with 5528 number the outcome is appears as in table3 which express that the Correctly Classified Instances 4540 which is in 82.1274 %, Incorrectly Classified Instances 988 which is 17.8726 %, Kappa measurement is 0.7855, Mean outright mistake is 0.07, Root mean squared blunder 0.187, Relative total blunder is 25.1833 %, Root relative squared blunder is 50.183 %.

Table3.1. Confusion Matrix of J48

a	b	c	d	e	f
879	1	26	0	17	0
0	746	0	77	1	107
26	0	886	0	4	0
0	310	0	556	3	52
35	6	7	1	862	5
4	266	1	26	13	611

Where, a = Labour force participation Total, b = Unemployment rate Total, c = Labour force participation Male, d = Unemployment rate Male, e = Labour force participation   Female, f = Unemployment rate Female.

Table4. Result of LMT

	No. Of instance	Result
Correctly Classified Instances	4384	79.3054%
Incorrectly Classified Instances	1144	20.6946%
Kappa statistic	0.7517	
Mean absolute error	0.0924	
Root mean squared error	0.2052	
Relative absolute error	33.249%	
Root relative squared error	55.0739%	
Total Number of Instances	5528	

Subsequent to playing out the LMT classifier activity on dataset with 5528 number the outcome is appears as in table4 which express that the Correctly Classified Instances 4384 which is in 79.3054 %, Incorrectly Classified Instances 1144

which is 20.6946 %, Kappa measurement is 0.7517, Mean supreme blunder is 0.0924, Root mean squared mistake 0.2052, Relative outright mistake is 33.249 %, Root relative squared mistake is 55.0739 %.

Table4.1. Confusion Matrix of LMT

a	b	c	d	e	f
841	0	28	1	53	0
1	492	0	232	2	204
10	0	900	0	6	0
0	113	0	661	1	146
104	0	14	1	786	11
5	99	0	105	8	704

Where, a = Labour force participation Total, b = Unemployment rate Total, c = Labour force participation Male, d = Unemployment rate Male, e = Labour force participation   Female, f = Unemployment rate Female.

Table5. Result of RandomForest

	No. Of instance	Result
Correctly Classified Instances	5445	98.4986%
Incorrectly Classified Instances	83	1.5014%
Kappa statistic	0.982	
Mean absolute error	0.0682	
Root mean squared error	0.1268	
Relative absolute error	24.5602%	
Root relative squared error	34.029%	
Total Number of Instances	5528	

Subsequent to playing out the RandomForest classifier activity on dataset with 5528 number the outcome is appears as in table 5 which express that the Correctly Classified Instances 5445 which is in 98.4986 %, Incorrectly Classified Instances 83 which is 1.5014 %, Kappa measurement is 0.982, Mean supreme mistake is 0.0682, Root mean squared blunder is 0.1268, Relative total mistake is 24.5602 %, Root relative squared mistake is 34.029 %.

Table5.1. Confusion Matrix of RandomForest

a	b	c	d	e	f
922	0	0	0	1	0
0	898	0	22	0	11
0	0	916	0	0	0
0	18	0	896	0	7
0	0	0	0	916	0
0	15	0	8	1	897

Where, a = Labour force participation Total, b = Unemployment rate Total, c = Labour force participation Male, d = Unemployment rate Male, e = Labour force participation   Female, f = Unemployment rate Female.

Table6. Result of RandomTree

	No. Of instance	Result
Correctly Classified Instances	5443	98.4624%
Incorrectly Classified Instances	85	1.5376%
Kappa statistic	0.9815	
Mean absolute error	0.0051	
Root mean squared error	0.0506	
Relative absolute error	1.8452%	
Root relative squared error	13.5837%	
Total Number of Instances	5528	

In the wake of playing out the RandomTree classifier movement on database with 5528 number the result is appeared as in table6 which express that the Correctly Classified Instances 5543 which is in 98.4624%, Incorrectly Classified Instances 83 which is 1.5376 %, Kappa estimation is 0.9815, Mean preeminent error is 0.0051, Root mean squared screw up 0.0506, Relative all out bungle is 1.8452%, Root relative squared bumble is 13.5837%.

Table6.1. Confusion Matrix RandomTree

a	b	c	d	e	f
923	0	0	0	0	0
0	931	0	0	0	0
0	0	916	0	0	0
0	49	0	872	0	0
1	0	0	0	915	0
0	34	0	0	1	886

Where, a = Labour force participation Total, b = Unemployment rate Total, c = Labour force participation Male, d = Unemployment rate Male, e = Labour force participation   Female, f = Unemployment rate Female.

Table7. Result of HoeffdingTree

	No. Of instance	Result
Correctly Classified Instances	3675	66.4797%
Incorrectly Classified Instances	1853	33.5203%
Kappa statistic	0.5976	
Mean absolute error	0.12	
Root mean squared error	0.2449	
Relative absolute error	43.1991%	
Root relative squared error	65.726%	
Total Number of Instances	5528	

In the wake of playing out the REPTree classifier activity on dataset with 5528 number the outcome is appears as in table7 which express that the Correctly Classified Instances 3675 which is in 66.4797%, Incorrectly Classified Instances 1853 which is 33.5203%, Kappa measurement is 0.5976, Mean total blunder is 0.12, Root mean squared mistake 0.2449, Relative outright blunder is 43.1991%, Root relative squared mistake is 65.726 %.

Table7.1. Confusion Matrix of REPTree

a	b	c	d	e	f
806	1	65	0	51	0
2	814	0	54	5	56
153	0	759	0	4	0
2	614	0	283	6	16
185	17	7	4	702	1
12	557	0	13	28	311

Where, a = Labour force participation Total, b = Unemployment rate Total, c = Labour force participation Male, d = Unemployment rate Male, e = Labour force participation   Female, f = Unemployment rate Female.

5   Conclusion

After the few classifier tries the result on the work power cooperation, Unemployment rate, in which Male support, and there joblessness rate,also the Female work power investment and there joblessness rate discovered exceptional outcome by using DecisionTree classifier, HoeffdingTree classifier, J48 classifier, LMT classifier, RandomForest classifier, RandomTree classifier, and in end REPTree classifier.When we apply DecisionTree classifier the outcome is poor strikingly with other classifier for instance 33.4298%. in addition, Other hand the RandomForest classifier result is high then again with other classifier for instance 98.4986%. On the off chance that in case we apply classifier of same kind information later on, we could phenomenal outcomes by the use of above classifier. Just DecisionTree classifier will be considered as phenomenal.

6   References

1. United Nations static division http://data.un.org
2. Geoff Hulten, Laurie Spencer, Pedro Domingos: "Mining time-changing data streams". In: ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining, 97-106, 2001.
3. Ross Quinlan (1993). C4.5: "Programs for Machine Learning". Morgan Kaufmann Publishers, San Mateo, CA.
4. Niels Landwehr, Mark Hall, Eibe Frank (2005). Logistic Model Trees. Machine Learning. 95(1-2):161-205George H. John, Pat Langley: "Estimating Contiguous Distributions in Bayesian Classifiers". In Eleventh Confrence on Uncertainly in Artificial Intelligence, San Mateo, 338-345,1995.
5. Leo Breiman (2001). Random Forests. Machine Learning. 45(1):5-32
6. ILO, Labour force projections, 1965-85 (1st edition, Geneva 1971).
7. ILO, Labour force projections, 1950-2010 (2nd edition, Geneva 1976).
8. ILO, Economically Active Population: Estimates and projections, 1950-2025 (3rd edition, Geneva 1986).
9. ILO, Economically Active Population Estimates and projections, 1950-2010 (4th edition, Geneva 1996).
10. ILO, Estimates and Projections of the Economically Active Population, 1980-2020 (5th edition, Geneva 2007, Update August 2008,Update December 2009).
11. The reduction from 193 to 189 reference areas is due to the inclusion of the data of Dominique, Martinique, French Guyana and Réunion into metropolitan France. The change in horizon from 2050 to 2030 allows to base the projections solely on econometricmethods.
12. ILO, Estimates and Projections of the Economically Active Population, 1990-2030 (2013 edition, Geneva 2013); ILO, Labour Force. Estimates and Projections, 1990-2050 (2015 edition, Geneva 2015).
13. Anupriya K. Kamble, Ramesh R. Manza, Yogesh M. Rajput, Kavita A. Khobragade "Classification of Insulin Dependent Diabetes Mellitus Blood Glucose Level Using Support Vector Machine", IOSR Journal of Computer Engineering (IOSR-JCE), e-ISSN: 2278-0661,p-ISSN: 2278-8727(2017).