

Analysis of Road Accidents through Multiple Mining Techniques

Rashmi H Patil¹, Siddu P Algur²

¹Department of Computer Science
Rani Chennamma University, Belagavi, India

²Department of Computer Science
Rani Chennamma University, Belagavi, India

ABSTRACT

Road accident is unpredictable and uncertain in entire world; India is the highest one among in the world. Road accident is one of main cause of property damage, unnatural deaths and disability, Road accident data analysis leads to improve the safety. In this paper we proposing two mining techniques decision tree and IF THEN rule to analyze the accident data. For this experiment we have 1,040,051 accident records from the year 2009 to 2015. Each record contains 30 factors; we have selected four major factors (Road type, Speed, Day of week and area). The decision tree analysis shows that the number of road accident occurred on Single carriageway is more (Average 2.69 %), road accident occurred in one-way street, Sunday & Saturday is less. By using this analysis can improve efficiently the level of Road accident safety management.

Keywords: Road accident; Decision tree, IF THEN; Analysis; Safety Management

1. INTRODUCTION

Road accident is unpredictable and uncertain in entire world, Road accident is one of main cause of property damage, unnatural deaths and disability. The frequency of traffic collisions in India is amongst the highest in the world. A national crime records Bureau (NCRB) report revealed that every year, more than 135,000 traffic collision-related deaths occur in India. Tamil Nadu records the highest road accidents for a decade and its capital Chennai has more accidents than any other city in India.

According to road traffic safety experts, the actual number of casualties may be higher than what is documented, as many traffic accidents go unreported. Moreover, victims who die some time after the accident, a span of time which may vary from a few hours to several days, are not counted as car accident victims [1].

When an accident took place, it is recorded by the concerned police officer of that region's police station. Road accidents contain the many data and information can analyze in multiple levels, angles using mining techniques.

Cafiso et. al., classify accidents according to their causes using a fuzzy algorithm through a computerized procedure and a simplified model [4]. Wongcharoen and Senivongse propose a method to predict traffic congestion severity level based on the analysis of Twitter messages. Different types of tweets data are used to construct a C4.5 decision tree model for prediction, including tweets from selected road-traffic Twitter accounts, tweets that contain road-traffic-related keywords, and geo-tagged tweets whose volume suggests large crowds in certain areas [5].

The rough sets theory and the theory of association rules are proposed to road traffic accidents causes' analysis [6]. The factor influencing the accidents and identified which factor is more accident prone with Info Gain Attribute Evaluator function using WEKA tool and also performed association classification using Apriori algorithm [7].

In this paper we proposing mining techniques like: Decision tree and IF THEN rules techniques to analyse road accidents based on day wise (Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday), area (Rural, Urban), speed of vehicle (less than 50km/h, greater than 50km/h) and types of road (Roundabout, One way street, Dual Carriageway, Single Carriageway, Slip road, Unknown) that accident occurred.

2. PROPOSED WORK

Road accident data analysis consist of two major mining methods they are: decision tree and IF THEN rules. The proposed system consists of two modules. In module one, we mainly analyzing four factors of road accident, those factors are showed in the Figure 1. Here we analyze number of road accidents occurred by speed, in day of week, Area, and type of roads.

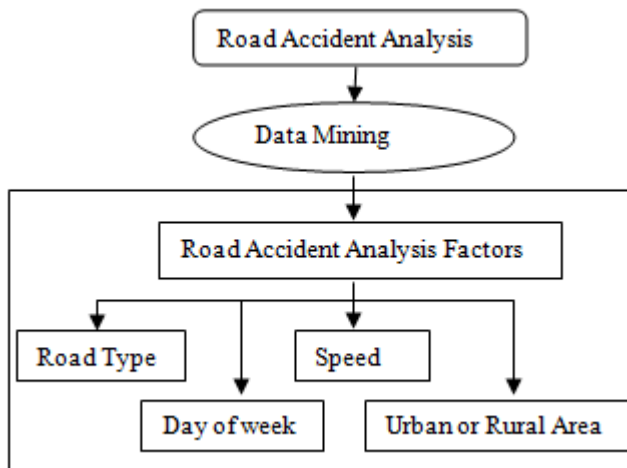


Figure1: Proposed Architecture

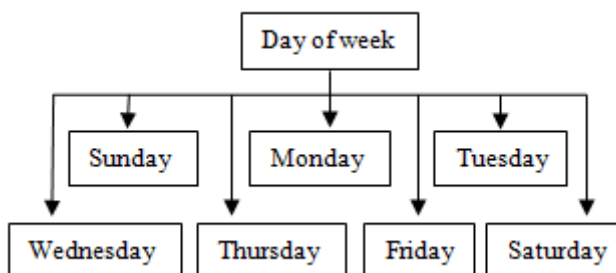


Figure 2: Day of Week Factors

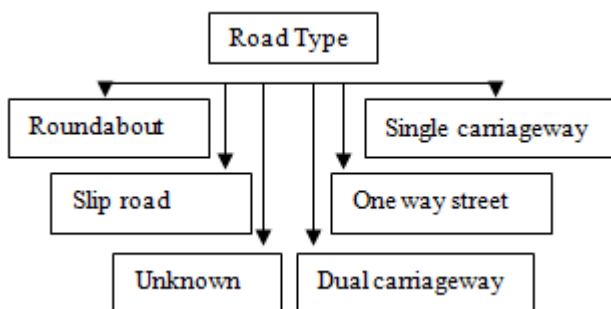


Figure 3: The types of road

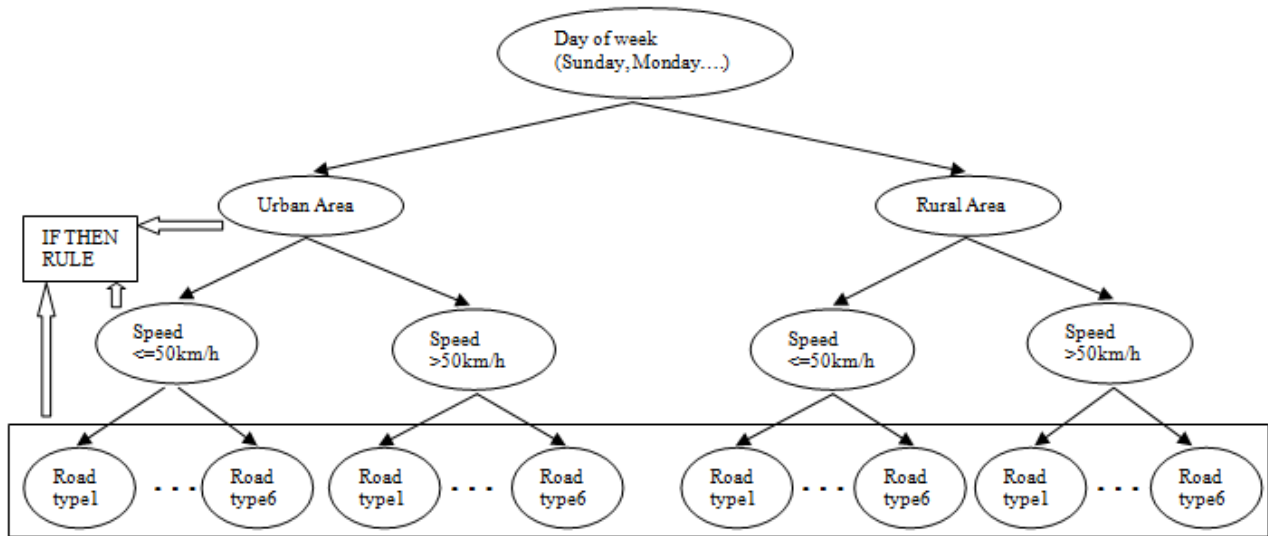


Figure 4: The Proposed Decision Tree

In module two, we analyzing inter relation of those four factors and also sub factors, we analyzing based on sub factors of day of week factors using Decision tree & IF THEN Rules, i.e. for example, the number of accidents occurred on one-way street with vehicle speed less than 50km/h in rural area on Sundays. The pseudo code for proposed system is showed below:

Pseudo-code: Road accident analysis

INPUT: Accident information includes 30 attributes that includes day of a week, area (Rural, Urban), speed of vehicle and types of road.

OUTPUT: Number of accident occurred on day wise with particular sub factors of area, speed and type of roads.

For all accidental records

IF accident occurred on particular day [exmp: Sunday] THEN

IF Area is Rural with speed <=50 THEN

Increase the count by one for respective type of road

IF Area is Rural with speed>50 THEN

Increase the count by one for respective type of road

IF Area is Urban with speed <=50 THEN

Increase the count by one for respective type of road

IF Area is Urban with speed>50 THEN

Increase the count by one for respective type of road

The Day of week factor contains 7 sub factors as shown in Figure 2. The type of roads factors are showed in Figure 3. The proposed decision tree is showed in Figure 4. By this Decision tree we can classify the number of accidents occurred in particular type of road. As shown in Figure 4, root node indicates the number of accident occurred on each day of week, let's consider root node is at level 1, then level two nodes indicate number of accident occurred in rural or urban area on each day of week, i.e., for example, Sunday how many accidents occurred on rural area, how many accidents are occurred on urban area, Likewise for all days of week. Level 3 nodes indicate day in weeks and in area and with speed less than or equal to 50km/h or greater than 50km/h. level 4 nodes indicate number of accidents occurred on type of the roads with all possibilities of above level.

3. EXPERIMENTAL RESULTS

We have road accidental data that fetch from UK database [3], it contains 7 years' [2009 to 2015] road accident information, there are totally 1,040,051 road accident records are noted, each year data contains at-least of one lack of accident reports and there are 30 factors in each report. In these we have selected four major factors; these four factors have sub-factors, types of road showed in Figure 3, Day of week in Figure 2, Area contain Urban and Rural, Factor Speed contain less than or equal to 50km/h and greater than 50km/h. From Table I to Table IV indicate the number of

accident occurred in day of week, Area, Speed and Types of road factors. Respective Average and Percentage is calculated. From Table V to Table X shows the number of accidents occurred for proposed decision tree, its average and percentage.

ANALYSIS FOR ACCIDENT OCCURRED ON DAY OF WEEK FACTOR

Day of Week Factors	years							Average	Percentage (%)
	2009	2010	2011	2012	2013	2014	2015		
Sunday	17861	16794	16274	15856	14854	15940	15258	16119.57	10.85
Monday	22972	22452	21392	20910	19831	21093	20032	21240.29	14.30
Tuesday	24791	23045	22470	21661	21257	22318	21431	22424.71	15.09
Wednesday	24799	23018	22926	21997	20461	22210	21368	22397.00	15.07
Thursday	24305	22810	23461	22334	21186	21780	21479	22479.29	15.13
Friday	26666	25475	24948	23511	23044	23960	22374	24282.57	16.34
Saturday	22160	20820	20003	19302	18027	19021	18114	19635.29	13.22

ANALYSIS FOR ROAD ACCIDENT OCCURRED ON AREA FACTOR

Area Factors	years							Average	Percentage
	2009	2010	2011	2012	2013	2014	2015		
Rural	59016	54926	52898	50862	48926	50035	47976	52091.29	35.06
Urban	104538	99488	98576	94709	89734	96287	92080	96487.43	64.94

ANALYSIS FOR ROAD ACCIDENT OCCURRED ON SPEED FACTOR

Speed Factors	years							Average	Percentage
	2009	2010	2011	2012	2013	2014	2015		
Speed <=50	125781	119596	118851	114377	108973	116229	111349	116450.86	78.38
Speed >50	37773	34818	32623	31194	29687	30093	28707	32127.86	21.62

ANALYSIS FOR ROAD ACCIDENT OCCURRED ON TYPE OF ROAD FACTOR

Type of Road Factors	years							Average	Percentage
	2009	2010	2011	2012	2013	2014	2015		
Roundabout	11041	10419	10514	10173	9405	10329	9589	10210.00	6.87
One way street	3271	3153	3049	2747	2744	2935	2873	2967.43	2.00
Dual carriageway	24052	22943	21474	20572	19751	20823	20266	21411.57	14.41
Single carriageway	122573	115532	114164	109970	104845	110187	105066	111762.43	75.22
Slip road	1698	1663	1544	1552	1423	1542	1456	1554.00	1.05
Unknown	919	704	729	557	492	506	806	673.29	0.45

DECISION TREE ANALYSIS FOR ROAD ACCIDENT OCCURRED ON TYPE OF ROAD FACTOR-ROUNDBABOUT (RA)

Factors	years							Average	Percentage
	2009	2010	2011	2012	2013	2014	2015		
Sunday-Urban-Less-RA	681	668	713	622	587	628	616	645.00	0.43
Sunday-Urban-Greater-RA	37	26	23	27	26	20	14	24.71	0.02
Sunday-Rural-Less-RA	229	223	239	238	219	284	245	239.57	0.16
Sunday-Rural- Greater-RA	240	200	182	208	164	179	205	196.86	0.13
Monday-Urban-Less-RA	978	973	958	951	890	953	889	941.71	0.63
Monday-Urban-Greater-RA	30	36	43	33	23	18	14	28.14	0.02
Monday-Rural-Less-RA	286	300	317	314	274	334	305	304.29	0.20
Monday-Rural- Greater-RA	260	226	215	199	170	213	185	209.71	0.14
Tuesday-Urban-Less-RA	1084	982	941	1003	944	1028	992	996.29	0.67
Tuesday-Urban-Greater-RA	47	36	39	38	30	24	20	33.43	0.02
Tuesday-Rural-Less-RA	323	314	335	329	312	366	340	331.29	0.22
Tuesday-Rural- Greater-RA	283	262	229	211	209	209	168	224.43	0.15
Wednesday-Urban-Less-RA	1085	1014	1065	971	945	1013	916	1001.29	0.67
Wednesday-Urban-Greater-RA	55	51	53	31	17	25	22	36.29	0.02
Wednesday-Rural-Less-RA	338	307	340	332	294	364	356	333.00	0.22
Wednesday-Rural- Greater-RA	241	239	222	227	203	183	186	214.43	0.14
Thursday-Urban-Less-RA	1064	937	1090	971	922	1015	969	995.43	0.67
Thursday-Urban-Greater-RA	49	46	36	24	36	21	20	33.14	0.02
Thursday-Rural-Less-RA	329	309	319	350	317	354	333	330.14	0.22
Thursday-Rural- Greater-RA	263	225	247	192	183	188	175	210.43	0.14

Friday-Urban-Less-RA	1077	1081	1056	1060	933	999	932	1019.71	0.69
Friday-Urban-Greater-RA	53	56	46	40	33	29	25	40.29	0.03
Friday-Rural-Less-RA	349	339	310	285	305	361	314	323.29	0.22
Friday-Rural- Greater-RA	243	267	247	242	226	227	177	232.71	0.16
Saturday-Urban-Less-RA	865	780	775	765	714	763	691	764.71	0.51
Saturday-Urban-Greater-RA	41	42	29	22	23	25	17	28.43	0.02
Saturday-Rural-Less-RA	287	272	248	283	246	334	292	280.29	0.19
Saturday-Rural- Greater-RA	224	208	197	205	160	172	171	191.00	0.13

DECISION TREE ANALYSIS FOR ROAD ACCIDENT OCCURRED ON TYPE OF ROAD FACTOR-ONE WAY STREET (OWS)

Factors	years							Average	Percentage
	2009	2010	2011	2012	2013	2014	2015		
Sunday-Urban-Less-OWS	311	265	273	251	239	274	294	272.43	0.18
Sunday-Urban-Greater-OWS	0	0	0	0	1	0	1	0.29	0.00
Sunday-Rural-Less-OWS	21	15	26	24	19	20	18	20.43	0.01
Sunday-Rural- Greater-OWS	2	6	4	5	4	4	3	4.00	0.00
Monday-Urban-Less-OWS	376	389	368	363	356	381	361	370.57	0.25
Monday-Urban-Greater-OWS	2	0	1	0	0	1	0	0.57	0.00
Monday-Rural-Less-OWS	25	35	35	31	24	22	30	28.86	0.02
Monday-Rural- Greater-OWS	6	4	4	2	6	2	3	3.86	0.00
Tuesday-Urban-Less-OWS	449	463	438	362	337	415	368	404.57	0.27
Tuesday-Urban-Greater-OWS	1	0	1	0	0	0	0	0.29	0.00
Tuesday-Rural-Less-OWS	34	39	40	23	35	25	27	31.86	0.02
Tuesday-Rural- Greater-OWS	5	9	5	5	3	2	3	4.57	0.00
Wednesday-Urban-Less-OWS	463	427	396	366	370	430	406	408.29	0.27
Wednesday-Urban-Greater-OWS	0	1	1	0	0	0	0	0.29	0.00
Wednesday-Rural-Less-OWS	30	27	37	21	27	22	27	27.29	0.02
Wednesday-Rural- Greater-OWS	4	6	5	3	7	6	4	5.00	0.00
Thursday-Urban-Less-OWS	439	437	362	353	373	402	406	396.00	0.27
Thursday-Urban-Greater-OWS	1	1	4	2	1	0	1	1.43	0.00
Thursday-Rural-Less-OWS	37	38	21	30	36	25	20	29.57	0.02
Thursday-Rural- Greater-OWS	4	3	7	8	5	2	5	4.86	0.00
Friday-Urban-Less-OWS	514	497	502	428	452	457	430	468.57	0.32
Friday-Urban-Greater-OWS	0	0	1	0	0	1	2	0.57	0.00
Friday-Rural-Less-OWS	43	46	39	39	31	30	28	36.57	0.02
Friday-Rural- Greater-OWS	8	11	9	3	6	4	3	6.29	0.00
Saturday-Urban-Less-OWS	455	408	436	397	386	383	399	409.14	0.28
Saturday-Urban-Greater-OWS	0	0	1	0	0	1	0	0.29	0.00
Saturday-Rural-Less-OWS	33	23	27	25	20	23	28	25.57	0.02
Saturday-Rural- Greater-OWS	8	3	6	6	6	3	6	5.43	0.00

DECISION TREE ANALYSIS FOR ROAD ACCIDENT OCCURRED ON TYPE OF ROAD FACTOR-DUAL CARRIAGEWAY (DC)

Factors	years							Average	Percentage
	2009	2010	2011	2012	2013	2014	2015		
Sunday-Urban-Less-DC	1102	1042	1002	963	924	997	942	996.00	0.67
Sunday-Urban-Greater-DC	144	159	123	123	121	141	143	136.29	0.09
Sunday-Rural-Less-DC	183	214	174	191	166	194	187	187.00	0.13
Sunday-Rural- Greater-DC	1247	1239	1116	1071	1018	1040	989	1102.86	0.74
Monday-Urban-Less-DC	1559	1542	1347	1347	1246	1332	1332	1386.43	0.93
Monday-Urban-Greater-DC	206	195	150	183	180	223	183	188.57	0.13
Monday-Rural-Less-DC	248	227	237	220	223	264	285	243.43	0.16
Monday-Rural- Greater-DC	1446	1418	1296	1215	1206	1221	1152	1279.14	0.86
Tuesday-Urban-Less-DC	1716	1544	1508	1471	1455	1392	1412	1499.71	1.01
Tuesday-Urban-Greater-DC	193	200	156	200	186	195	224	193.43	0.13
Tuesday-Rural-Less-DC	264	275	236	233	225	279	337	264.14	0.18
Tuesday-Rural- Greater-DC	1323	1307	1222	1144	1197	1267	1172	1233.14	0.83
Wednesday-Urban-Less-DC	1728	1522	1562	1428	1287	1440	1333	1471.43	0.99
Wednesday-Urban-Greater-DC	216	201	188	188	180	209	208	198.57	0.13
Wednesday-Rural-Less-DC	266	251	236	238	201	291	276	251.29	0.17
Wednesday-Rural- Greater-DC	1410	1341	1294	1251	1102	1196	1162	1250.86	0.84
Thursday-Urban-Less-DC	1689	1567	1548	1494	1345	1361	1405	1487.00	1.00
Thursday-Urban-Greater-DC	226	190	177	208	166	218	199	197.71	0.13
Thursday-Rural-Less-DC	262	262	266	237	247	250	275	257.00	0.17
Thursday-Rural- Greater-DC	1416	1344	1319	1224	1206	1142	1161	1258.86	0.85

Friday-Urban-Less-DC	1800	1723	1668	1526	1460	1535	1425	1591.00	1.07
Friday-Urban-Greater-DC	252	213	198	186	180	244	210	211.86	0.14
Friday-Rural-Less-DC	274	279	248	226	244	267	321	265.57	0.18
Friday-Rural- Greater-DC	1719	1660	1467	1431	1415	1468	1355	1502.14	1.01
Saturday-Urban-Less-DC	1391	1332	1338	1240	1160	1247	1214	1274.57	0.86
Saturday-Urban-Greater-DC	191	160	127	108	136	161	166	149.86	0.10
Saturday-Rural-Less-DC	227	246	199	188	237	203	239	219.86	0.15
Saturday-Rural- Greater-DC	1354	1290	1072	1038	1038	1046	959	1113.86	0.75

DECISION TREE ANALYSIS FOR ROAD ACCIDENT OCCURRED ON TYPE OF ROAD FACTOR-SINGLE CARRIAGEWAY (SC)

Factors	years							Average	Percent age
	2009	2010	2011	2012	2013	2014	2015		
Sunday-Urban-Less-SC	7952	7511	7178	7247	6755	7400	7080	7303.29	4.92
Sunday-Urban-Greater-SC	72	54	68	60	43	44	42	54.71	0.04
Sunday-Rural-Less-SC	2239	2107	2209	2063	2062	2137	1979	2113.71	1.42
Sunday-Rural- Greater-SC	3111	2803	2707	2536	2325	2334	2222	2576.86	1.73
Monday-Urban-Less-SC	11321	11227	10956	10562	9916	10834	10120	10705.14	7.21
Monday-Urban-Greater-SC	100	82	70	64	67	52	71	72.29	0.05
Monday-Rural-Less-SC	2711	2614	2454	2519	2472	2480	2351	2514.43	1.69
Monday-Rural- Greater-SC	3041	2864	2576	2594	2486	2435	2406	2628.86	1.77
Tuesday-Urban-Less-SC	12382	11663	11640	11207	10932	11586	11127	11505.29	7.74
Tuesday-Urban-Greater-SC	114	82	83	77	80	60	51	78.14	0.05
Tuesday-Rural-Less-SC	2929	2635	2626	2555	2597	2636	2556	2647.71	1.78
Tuesday-Rural- Greater-SC	3211	2847	2650	2492	2401	2541	2315	2636.71	1.77
Wednesday-Urban-Less-SC	12268	11762	11668	11275	10550	11695	11093	11473.00	7.72
Wednesday-Urban-Greater-SC	108	74	88	82	92	74	83	85.86	0.06
Wednesday-Rural-Less-SC	2897	2644	2688	2622	2463	2515	2507	2619.43	1.76
Wednesday-Rural- Greater-SC	3308	2796	2740	2644	2438	2449	2435	2687.14	1.81
Thursday-Urban-Less-SC	12204	11674	11963	11608	10973	11389	11299	11587.14	7.80
Thursday-Urban-Greater-SC	95	80	90	75	75	69	69	79.00	0.05
Thursday-Rural-Less-SC	2741	2547	2770	2555	2557	2573	2483	2603.71	1.75
Thursday-Rural- Greater-SC	3069	2796	2867	2688	2454	2455	2350	2668.43	1.80
Friday-Urban-Less-SC	13328	12717	12886	11904	11696	12519	11549	12371.29	8.33
Friday-Urban-Greater-SC	113	98	86	75	71	66	76	83.57	0.06
Friday-Rural-Less-SC	2999	2942	2879	2874	2859	2730	2617	2842.86	1.91
Friday-Rural- Greater-SC	3510	3142	2954	2833	2839	2700	2543	2931.57	1.97
Saturday-Urban-Less-SC	10545	10086	9929	9716	8856	9497	9048	9668.14	6.51
Saturday-Urban-Greater-SC	82	73	69	49	67	65	39	63.43	0.04
Saturday-Rural-Less-SC	2794	2508	2463	2360	2216	2354	2176	2410.14	1.62
Saturday-Rural- Greater-SC	3329	3104	2807	2634	2503	2498	2379	2750.57	1.85

DECISION TREE ANALYSIS FOR ROAD ACCIDENT OCCURRED ON TYPE OF ROAD FACTOR-SLIP ROAD (SR)

Factors	years							Average	Percentage
	2009	2010	2011	2012	2013	2014	2015		
Sunday-Urban-Less-SR	56	46	48	54	42	47	52	49.29	0.03
Sunday-Urban-Greater-SR	12	14	14	14	11	21	20	15.14	0.01
Sunday-Rural-Less-SR	22	19	15	18	9	30	24	19.57	0.01
Sunday-Rural- Greater-SR	111	101	100	97	76	90	82	93.86	0.06
Monday-Urban-Less-SR	79	76	77	70	80	74	73	75.57	0.05
Monday-Urban-Greater-SR	16	23	14	15	13	23	21	17.86	0.01
Monday-Rural-Less-SR	38	28	34	31	28	26	31	30.86	0.02
Monday-Rural- Greater-SR	113	104	116	113	88	122	114	110.00	0.07
Tuesday-Urban-Less-SR	85	84	69	62	64	65	63	70.29	0.05
Tuesday-Urban-Greater-SR	15	19	16	15	31	15	14	17.86	0.01
Tuesday-Rural-Less-SR	40	38	28	24	30	33	31	32.00	0.02
Tuesday-Rural- Greater-SR	134	141	105	127	103	108	90	115.43	0.08
Wednesday-Urban-Less-SR	84	81	70	66	74	74	68	73.86	0.05
Wednesday-Urban-Greater-SR	17	11	20	20	15	24	22	18.43	0.01
Wednesday-Rural-Less-SR	31	40	34	32	30	36	34	33.86	0.02
Wednesday-Rural- Greater-SR	121	129	101	113	110	91	99	109.14	0.07
Thursday-Urban-Less-SR	88	87	69	60	68	78	56	72.29	0.05
Thursday-Urban-Greater-SR	26	16	14	23	18	19	12	18.29	0.01
Thursday-Rural-Less-SR	27	25	34	32	26	36	25	29.29	0.02
Thursday-Rural- Greater-SR	124	103	137	112	93	111	101	111.57	0.08
Friday-Urban-Less-SR	84	81	75	80	76	87	71	79.14	0.05

Friday-Urban-Greater-SR	14	21	25	25	23	14	15	19.57	0.01
Friday-Rural-Less-SR	37	45	31	37	20	30	39	34.14	0.02
Friday-Rural- Greater-SR	118	141	121	119	104	110	104	116.71	0.08
Saturday-Urban-Less-SR	54	59	52	58	58	48	62	55.86	0.04
Saturday-Urban-Greater-SR	21	17	13	16	19	18	17	17.29	0.01
Saturday-Rural-Less-SR	33	29	20	19	19	30	34	26.29	0.02
Saturday-Rural- Greater-SR	98	85	92	100	95	82	82	90.57	0.06

DECISION TREE ANALYSIS FOR ROAD ACCIDENT OCCURRED ON TYPE OF ROAD FACTOR-UNKNOWN (UK)

Factors	years							Average	Percentage
	2009	2010	2011	2012	2013	2014	2015		
Sunday-Urban-Less-UK	60	45	35	29	28	37	67	43.00	0.03
Sunday-Urban-Greater-UK	1	1	1	0	0	1	1	0.71	0.00
Sunday-Rural- Less -UK	18	18	12	9	8	10	20	13.57	0.01
Sunday-Rural- Greater -UK	10	18	12	6	7	8	12	10.43	0.01
Monday-Urban-Less-UK	89	64	79	56	54	54	80	68.00	0.05
Monday-Urban-Greater-UK	0	0	1	0	1	0	2	0.57	0.00
Monday-Rural- Less -UK	30	15	24	14	18	19	11	18.71	0.01
Monday-Rural- Greater -UK	12	10	20	14	10	10	13	12.71	0.01
Tuesday-Urban-Less-UK	116	74	68	58	49	49	80	70.57	0.05
Tuesday-Urban-Greater-UK	3	1	2	0	1	0	0	1.00	0.00
Tuesday-Rural- Less -UK	26	17	19	15	22	14	22	19.29	0.01
Tuesday-Rural- Greater -UK	14	13	14	10	14	9	19	13.29	0.01
Wednesday-Urban-Less-UK	82	63	73	58	35	50	93	64.86	0.04
Wednesday-Urban-Greater-UK	2	0	2	0	0	1	1	0.86	0.00
Wednesday-Rural- Less -UK	30	16	20	17	10	15	23	18.71	0.01
Wednesday-Rural- Greater -UK	15	15	23	12	11	7	14	13.86	0.01
Thursday-Urban-Less-UK	105	83	87	63	50	44	79	73.00	0.05
Thursday-Urban-Greater-UK	1	4	1	1	0	0	1	1.14	0.00
Thursday-Rural- Less -UK	29	25	20	12	26	16	20	21.14	0.01
Thursday-Rural- Greater -UK	17	11	13	12	9	12	15	12.71	0.01
Friday-Urban-Less-UK	94	65	66	69	49	55	101	71.29	0.05
Friday-Urban-Greater-UK	1	2	0	1	0	1	1	0.86	0.00
Friday-Rural- Less -UK	16	26	15	15	11	14	16	16.14	0.01
Friday-Rural- Greater -UK	20	23	19	13	11	12	20	16.86	0.01
Saturday-Urban-Less-UK	79	57	67	50	38	37	59	55.29	0.04
Saturday-Urban-Greater-UK	0	5	1	0	0	0	1	1.00	0.00
Saturday-Rural- Less -UK	25	19	26	13	20	19	15	19.57	0.01
Saturday-Rural- Greater-UK	24	14	9	10	10	12	20	14.14	0.01

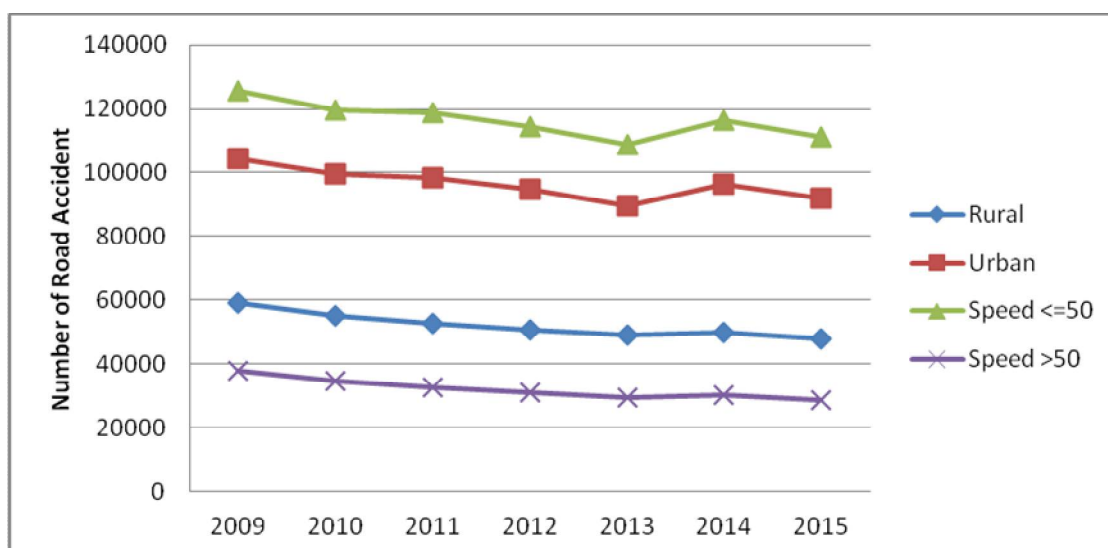


Figure 5: Curve with number of road accident for area & speed factors

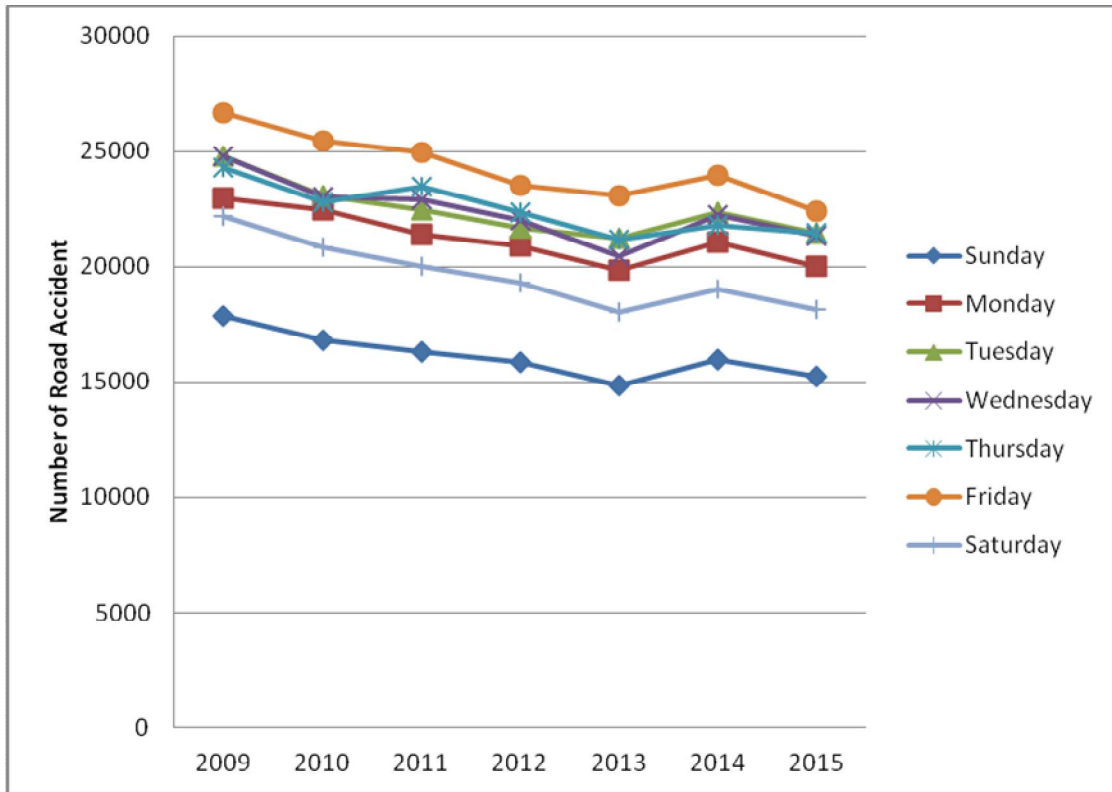


Figure 6: Curve with number of road accident for day of week factor

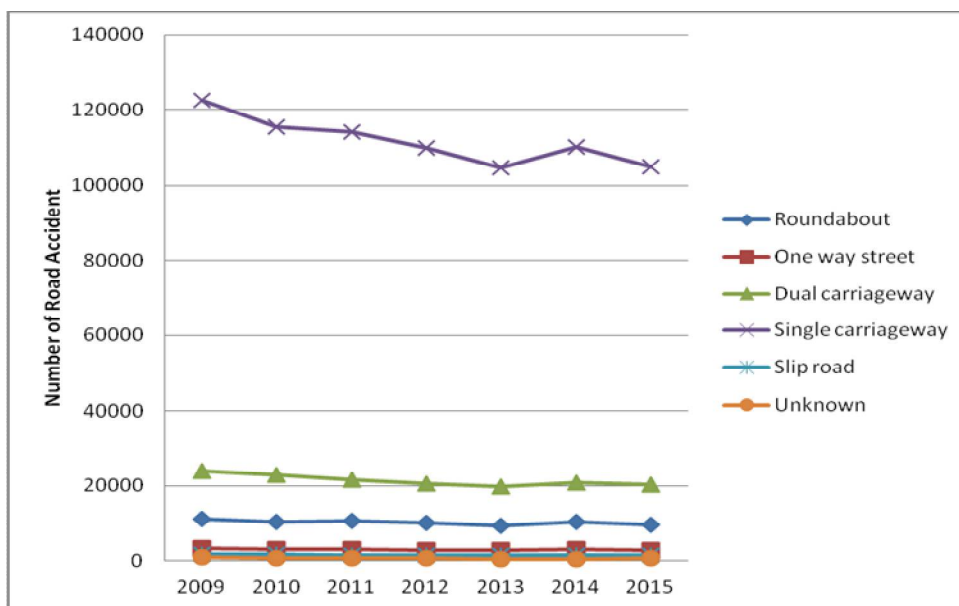


Figure 7: Curve with number of road accident for types of road factor

Figure 5, Figure 6 and Figure 7 shows the number of road accident analysis for Area-Speed, Day of Week and Type of roads respectively. In Figure 6, the number of road accident happen on Sunday & Saturday is lesser. By noticing Figure 7, the number of accident happen in single carriage way is more than any type of road. We can use these analysis data to predict and can improve the Road accidental safety management effectively.

4. CONCLUSION

In this paper we proposed two mining techniques decision tree and IF THEN rule to analyze the Road accident data. For this experiment we have 1,040,051 accident records from the year 2009 to till 2015. Each record contains 30 factors; we have selected four major factors (Road type, Speed, Day of week and area). The experimental result shows that, the number of road accident occurred on Single carriageway is more in term of percentage 2.69 (Average), road accident occurred in one-way street, Sunday & Saturday is less. Further can consider few more factors to analyze, can classify speed of vehicle factor to lesser than 30km/h, in between 30 to 50km/h and greater than 50km/h, we can use these analysis data to predict and can improve the Road accidental safety management effectively.

References

- [1] Murali Krishnan (29 April 2010). "India has the highest number of road accidents in the world". Deutsche Welle. Retrieved 3 May 2012.
- [2] <http://morth.nic.in/showfile.asp?lid=2143>
- [3] <https://data.gov.uk/dataset/road-accidents-safety-data>
- [4] Salvatore Cafiso, Grazia La Cava, Vincenzo Cutello, "A Fuzzy Model for Road Accidents Analysis", Fuzzy Information Processing Society, NAFIPS, 18th International Conference of the North American, 1999.
- [5] Sakkachin Wongcharoen and Twittie Senivongse, "Twitter Analysis of Road Traffic Congestion Severity Estimation", 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2016.
- [6] Rui Tian and Zhaosheng Yang, Maolei Zhang, "Method of Road Traffic Accidents Causes Analysis Based on Data Mining", International Conference on Computational Intelligence and Software Engineering (CiSE), 2010.
- [7] Suwarna Gothane, Dr. M. V. Sarode, "Analyzing Factors, Construction of Dataset, Estimating importance of factor and generation of association rules for Indian road Accident", 6th International Conference on Advanced Computing, IEEE 2016.
- [8] LI Jian, FAN Xiao-jun, HUANG Pei. Theory of Knowledge Based on Rough Sets and Its Application[J]. SYSTEMS ENGINEERING THEORY METHODOLOGY APPLICATIONS.2001,Vol.10 No.3, pp.184-188.a
- [9] He Ming. The System Research on The Traffic Accident Data Analysis Based on Data Mining[D]. Changsha University of Science & Technology, 2009.3.
- [10] Hümeyra Bolakar and Ahmet Tortum, "Clustering of Districts in Erzurum by Number of Injury", Journal of Traffic and Logistics Engineering vol. 3, pp. 125–128, Dec. 2015.

AUTHORS



Dr. Siddu P. Algur is working as Registrar and Professor in Dept of Computer Science, Rani Channamma University, Belagavi, Karnataka, India. He received his B.E. degree in Electrical and Electronics from Mysore University, Karnataka, India, in 1986. He received his M.E. degree from NIT, Allahabad, India, in 1991. He obtained Ph.D. degree from the Department of P.G. Studies and Research in Computer Science at Gulbarga University, Gulbarga. His research interest includes Data Mining, Web Mining, Big Data and Information Retrieval from the web and Knowledge discovery techniques. He has published more than 45 research papers in peer reviewed International Journals and chaired the sessions in many International conferences.



Mrs. Rashmi H Patil is pursuing her Ph.D programme in Computer Science, Rani Channamma University Belagavi, Karnataka, India. She received her BCA degree from Karnataka University, Dharwad, Karnataka, India, in 2009 and She received M.Sc in Computer Science degree from Rani Channamma University, Belgavi, Karnataka, India, in 2012. Her research interest includes Data Mining, Web Mining, Information Retrieval from the web and Knowledge discovery techniques. She published her research papers in International Journals and conferences.