Predicting Vehicle Crash Type from Crash Report Incident

Gabriel Xu, Andrew Chen

Problem and Background

- Vehicle crashes are common, over **6 million** per year in the U.S.
- Resulted in **40,990 deaths** in the U.S. in 2023
- In order to prevent this, consider which factors influence the **severity of crashes**
  - Allows us to know how better to **prevent injuries** and **fatalities**
  - Identify **which locations** are more **dangerous**

Dataset Overview

- 24241 instances
- 42 attributes
- Vehicle Crashes occurring in the Town of Cary, North Carolina
- Classifying instances as “fatalities”, “injury”, or “crash”
  - Required preprocessing steps
- Examples of attributes:
  - rdsurface
  - rdcondition
  - lightcond
  - weather
  - trafcontrol

Preprocessing

- Removed columns
  - tamainid
  - records
  - contributing_factor
  - vehicle_type
  - contrcir1_desc, contrcir2_desc, ...
  - vehicle3, vehicle4, ...
- New/modified columns
  - num_vehicles
  - class
  - ta_time
- Normalize data
- 24241 instances, 20 columns
- 0.18% fatalities, 13.34% injury, 86.47% crash

Dataset after Preprocessing

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	rdfeature	rdcharacter	rdclass	rdconfigur	rdsurface	rdcondition	lightcond	weather	trafcontrl	lon2	lat2	zone	numpassengers	numpedestrians	vehicle1	vehicle2	workarea	ta_time	month	vehicles	class
2	NO SPECIAL FEATU	STRAIGHT,LEVEL'	STATE SECONDAR'	TWO-WAY,DIVIDED,	COARSE ASPHALT'	DRY	DAYLIGHT	CLEAR	STOP SIGN'	-0.061311	-0.014066	WAKE	0	-0.078324	PASSENGER CAR'	SPORT UTILITY'	NO	-0.333225	-1.605539	0.288377	crash
3	NO SPECIAL FEATU	STRAIGHT,LEVEL'	PUBLIC VEHICULAR	ONE-WAY,NOT DIVI	COARSE ASPHALT'	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0.015774	-0.014067	112	-0.635271	-0.078324	PASSENGER CAR'	PASSENGER CAR'	NO	-0.128985	-1.605539	0.288377	crash
4	NO SPECIAL FEATU	STRAIGHT,LEVEL'	PUBLIC VEHICULAR	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0.015774	-0.014066	112	0	-0.078324	PICKUP	PASSENGER CAR'	NO	-0.537466	-1.605539	0.288377	crash
5	NO SPECIAL FEATU	STRAIGHT,GRADE'	PUBLIC VEHICULAR	ONE-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0.015774	-0.014067	112	0	-0.078324	PASSENGER CAR'	PASSENGER CAR'	NO	-1.354428	-1.313731	0.288377	crash
6	NO SPECIAL FEATU	STRAIGHT,LEVEL'	LOCAL STREET'	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DARK - ROADWAY I	CLEAR	NO CONTROL PRES	0.015774	-0.014066	122	-0.635271	-0.078324	FIRETRUCK	PASSENGER CAR'	NO	1.300699	-1.313731	-1.476022	crash
7	T-INTERSECTION	STRAIGHT,LEVEL'	LOCAL STREET'	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DAYLIGHT	CLEAR	STOP AND GO SIGN	0.015774	-0.014066	120	0	-0.078324	SPORT UTILITY'	PASSENGER CAR'	NO	-1.354428	-1.313731	0.288377	crash
8	NO SPECIAL FEATU	STRAIGHT,LEVEL'	STATE SECONDAR'	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0.015774	-0.014066	118	0	-0.078324	PASSENGER CAR'	PASSENGER CAR'	NO	-0.945947	-1.313731	-1.476022	crash
9	RELATED TO INTEF	STRAIGHT,LEVEL'	US ROUTE'	TWO-WAY,DIVIDED,	SMOOTH ASPHALT	DRY	DAYLIGHT	CLEAR	STOP AND GO SIGN	0.015774	-0.014067	112	0	-0.078324	SPORT UTILITY'	PASSENGER CAR'	NO	-0.945947	-1.313731	0.288377	injury
10	NO SPECIAL FEATU	STRAIGHT,LEVEL'	STATE SECONDAR'	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DAYLIGHT	CLEAR	STOP SIGN'	0.015774	-0.014066	112	0	-0.078324	VAN	PASSENGER CAR'	NO	0.483737	-1.605539	0.288377	injury
11	NO SPECIAL FEATU	STRAIGHT,LEVEL'	LOCAL STREET'	TWO-WAY,NOT DIVI	COARSE ASPHALT'	DRY	DAYLIGHT	CLOUDY	NO CONTROL PRES	0.015774	-0.014066	112	0.18559	-0.078324	PASSENGER CAR'	SPORT UTILITY'	NO	0.483737	-1.605539	0.288377	crash
12	NO SPECIAL FEATU	STRAIGHT,LEVEL'	PUBLIC VEHICULAR	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DAYLIGHT	CLOUDY	NO CONTROL PRES	0.015774	-0.014066	112	0	-0.078324	PASSENGER CAR'	SPORT UTILITY'	NO	0.892218	-1.313731	0.288377	crash
13	NO SPECIAL FEATU	CURVE,LEVEL'	STATE SECONDAR'	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0.015774	-0.014066	112	-0.635271	-0.078324	TRUCK/TRAILER	PASSENGER CAR'	NO	0.279496	-1.313731	-1.476022	crash
14	NO SPECIAL FEATU	STRAIGHT,LEVEL'	US ROUTE'	TWO-WAY,DIVIDED,	SMOOTH ASPHALT	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0.015774	-0.014067	116	0	-0.078324	PASSENGER CAR'	PASSENGER CAR'	NO	-1.150188	-1.313731	-1.476022	crash
15	DRIVEWAY, PUBLIC	STRAIGHT,LEVEL'	PUBLIC VEHICULAR	TWO-WAY,NOT DIVI	COARSE ASPHALT'	WET	DARK - LIGHTED RC	CLOUDY	NO CONTROL PRES	0.015928	-0.014066	112	0	-0.078324	PASSENGER CAR'	PASSENGER CAR'	NO	1.300699	-1.313731	0.288377	crash
16	NO SPECIAL FEATU	STRAIGHT,LEVEL'	LOCAL STREET'	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	WET	DARK - LIGHTED RC	CLOUDY	STOP SIGN'	0.015774	-0.014066	112	0	-0.078324	SPORT UTILITY'	PASSENGER CAR'	NO	1.709181	-1.313731	-1.476022	crash
17	NO SPECIAL FEATU	STRAIGHT,LEVEL'	PUBLIC VEHICULAR	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0.015774	-0.014066	112	0	-0.078324	PASSENGER CAR'	SPORT UTILITY'	NO	0.075256	-1.313731	0.288377	crash
18	NO SPECIAL FEATU	STRAIGHT,LEVEL'	PUBLIC VEHICULAR	TWO-WAY,NOT DIVI	COARSE ASPHALT'	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0.015774	-0.014066	RALW	0.18559	-0.078324	PASSENGER CAR'	PASSENGER CAR'	NO	0.483737	-1.313731	0.288377	crash
19	NO SPECIAL FEATU	STRAIGHT,LEVEL'	LOCAL STREET'	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0.015774	-0.014066	120	0	-0.078324	SPORT UTILITY'	PASSENGER CAR'	NO	0.483737	-1.313731	0.288377	crash
20	NO SPECIAL FEATU	STRAIGHT,LEVEL'	PUBLIC VEHICULAR	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DAYLIGHT	CLEAR	STOP SIGN'	0.015774	-0.014066	112	1.006451	-0.078324	PASSENGER CAR'	PICKUP	NO	-0.741706	-1.313731	0.288377	crash
21	NO SPECIAL FEATU	STRAIGHT,LEVEL'	LOCAL STREET'	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DARK - LIGHTED RC	CLEAR	NO CONTROL PRES	0.015774	-0.014066	117	0	-0.078324	SINGLE UNIT TRUC	PASSENGER CAR'	NO	-2.375631	-1.313731	0.288377	crash
22	NO SPECIAL FEATU	STRAIGHT,LEVEL'	PUBLIC VEHICULAR	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DAYLIGHT	CLOUDY	NO CONTROL PRES	0.015774	-0.014066	112	0	-0.078324	SPORT UTILITY'	PASSENGER CAR'	NO	0.483737	-1.313731	0.288377	crash
23	FOUR-WAY INTERSI	STRAIGHT,LEVEL'	STATE SECONDAR'	TWO-WAY,DIVIDED,	SMOOTH ASPHALT	DRY	DARK - LIGHTED RC	CLEAR	STOP AND GO SIGN	0.015774	-0.014066	115	0	-0.078324	PASSENGER CAR'	PASSENGER CAR'	NO	0.892218	-1.313731	0.288377	crash
24	NO SPECIAL FEATU	STRAIGHT,LEVEL'	PRIVATE ROAD,DRI	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	WET	DAYLIGHT	RAIN	NO CONTROL PRES	0.015774	-0.014066	112	0	-0.078324	PASSENGER CAR'	PICKUP	NO	0.075256	-1.313731	0.288377	crash
25	NO SPECIAL FEATU	STRAIGHT,LEVEL'	LOCAL STREET'	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DARK - LIGHTED RC	CLEAR	NO CONTROL PRES	0.015774	-0.014066	117	0	-0.078324	LIGHT TRUCK (MINI	PASSENGER CAR'	NO	1.300699	-1.313731	-1.476022	crash
26	NO SPECIAL FEATU	STRAIGHT,LEVEL'	STATE SECONDAR'	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	WET	DAYLIGHT	CLOUDY	NO CONTROL PRES	0.015774	-0.014067	116	0	-0.078324	SPORT UTILITY'	PICKUP	NO	-0.741706	-1.313731	0.288377	crash
27	NO SPECIAL FEATU	STRAIGHT,LEVEL'	US ROUTE'	TWO-WAY,DIVIDED,	COARSE ASPHALT'	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0.015774	-0.014066	116	0	-0.078324	PASSENGER CAR'	PASSENGER CAR'	NO	0.687978	-1.313731	-1.476022	crash
28	NO SPECIAL FEATU	CURVE,GRADE'	STATE SECONDAR'	TWO-WAY,DIVIDED,	COARSE ASPHALT'	DRY	DARK - LIGHTED RC	CLEAR	NO CONTROL PRES	0.015774	-0.014067	112	0	-0.078324	PASSENGER CAR'	PASSENGER CAR'	NO	-2.579872	-1.313731	0.288377	crash
29	NO SPECIAL FEATU	STRAIGHT,LEVEL'	PUBLIC VEHICULAR	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DARK - LIGHTED RC	CLEAR	NO CONTROL PRES	0.015774	-0.014066	112	0	-0.078324	PICKUP	VAN	NO	1.300699	-1.313731	0.288377	crash
30	DRIVEWAY, PUBLIC	STRAIGHT,LEVEL'	PUBLIC VEHICULAR	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DARK - LIGHTED RC	CLEAR	NO CONTROL PRES	0.015774	-0.014066	112	0	-0.078324	SPORT UTILITY'	PASSENGER CAR'	NO	-2.579872	-1.313731	0.288377	crash
31	FOUR-WAY INTERSI	STRAIGHT,LEVEL'	STATE SECONDAR'	TWO-WAY,DIVIDED,	SMOOTH ASPHALT	DRY	DARK - LIGHTED RC	CLEAR	NO CONTROL PRES	0.015774	-0.014066	112	0	-0.078324	SPORT UTILITY'	SPORT UTILITY'	NO	1.096459	-1.313731	0.288377	injury
32	NO SPECIAL FEATU	STRAIGHT,LEVEL'	PUBLIC VEHICULAR	TWO-WAY,DIVIDED,	COARSE ASPHALT'	DRY	DAYLIGHT	CLEAR	STOP SIGN'	0.015774	-0.014066	WAKE	0	-0.078324	PASSENGER CAR'	SPORT UTILITY'	NO	-1.354428	-1.313731	0.288377	crash
33	NO SPECIAL FEATU	STRAIGHT,GRADE'	LOCAL STREET'	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0	0	122	14.961093	-0.078324	PASSENGER CAR'	PASSENGER CAR'	NO	0.279496	-1.313731	-1.476022	crash
34	DRIVEWAY, PRIVAT	STRAIGHT,LEVEL'	PRIVATE ROAD,DRI	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0.015774	-0.014066	112	0	-0.078324	SINGLE UNIT TRUC	PASSENGER CAR'	NO	-0.333225	-1.313731	-1.476022	crash
35	T-INTERSECTION	STRAIGHT,LEVEL'	US ROUTE'	TWO-WAY,DIVIDED,	SMOOTH ASPHALT	DRY	DAYLIGHT	CLEAR	STOP AND GO SIGN	0.015774	-0.014067	116	0	-0.078324	PASSENGER CAR'	SPORT UTILITY'	NO	0.687978	-1.313731	2.052776	injury
36	FOUR-WAY INTERSI	STRAIGHT,HILLCRE	STATE SECONDAR'	TWO-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DARK - LIGHTED RC	CLEAR	STOP AND GO SIGN	0.015774	-0.014066	112	3.469035	-0.078324	SPORT UTILITY'	SPORT UTILITY'	NO	-0.741706	-1.313731	0.288377	injury
37	NO SPECIAL FEATU	STRAIGHT,LEVEL'	LOCAL STREET'	TWO-WAY,DIVIDED,	SMOOTH ASPHALT	DRY	DARK - LIGHTED RC	CLEAR	NO CONTROL PRES	0.015774	-0.014066	111	0	-0.078324	SPORT UTILITY'	PASSENGER CAR'	NO	-0.945947	-1.313731	-1.476022	crash
38	NO SPECIAL FEATU	STRAIGHT,LEVEL'	PUBLIC VEHICULAR	TWO-WAY,NOT DIVI	COARSE ASPHALT'	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0.015774	-0.014066	RALW	0	-0.078324	PASSENGER CAR'	PASSENGER CAR'	NO	0.279496	-1.313731	-1.476022	crash
39	RELATED TO INTEF	STRAIGHT,LEVEL'	STATE SECONDAR'	TWO-WAY,DIVIDED,	COARSE ASPHALT'	DRY	DARK - LIGHTED RC	CLEAR	STOP AND GO SIGN	0.015928	-0.014066	112	0	-0.078324	PASSENGER CAR'	SPORT UTILITY'	NO	0.687978	-1.313731	0.288377	crash
40	NO SPECIAL FEATU	STRAIGHT,LEVEL'	US ROUTE'	TWO-WAY,DIVIDED,	CONCRETE	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0.015774	-0.014066	112	-0.635271	-0.078324	PASSENGER CAR'	PASSENGER CAR'	NO	0.483737	-1.313731	3.817175	crash
41	NO SPECIAL FEATU	STRAIGHT,LEVEL'	US ROUTE'	TWO-WAY,DIVIDED,	GROOVED CONCRE	WET	DARK - LIGHTED RC	RAIN	NO CONTROL PRES	0.015774	-0.014067	112	0	-0.078324	PASSENGER CAR'	PASSENGER CAR'	NO	-1.96715	-1.313731	-1.476022	crash
42	NO SPECIAL FEATU	STRAIGHT,LEVEL'	STATE SECONDAR'	TWO-WAY,DIVIDED,	SMOOTH ASPHALT	WET	DAYLIGHT	CLEAR	STOP AND GO SIGN	0.015774	-0.014067	118	0	-0.078324	VAN	PASSENGER CAR'	NO	-1.354428	-1.313731	0.288377	crash
43	ON-RAMP ENTRY'	CURVE,GRADE'	US ROUTE'	ONE-WAY,NOT DIVI	SMOOTH ASPHALT	DRY	DARK - LIGHTED RC	CLEAR	NO CONTROL PRES	0.015774	-0.014066	112	0	-0.078324	SPORT UTILITY'	SPORT UTILITY'	NO	0.892218	-1.313731	0.288377	crash
44	NO SPECIAL FEATU	STRAIGHT,LEVEL'	PUBLIC VEHICULAR	TWO-WAY,NOT DIVI	COARSE ASPHALT'	DRY	DARK - LIGHTED RC	CLEAR	NO CONTROL PRES	0.015774	-0.014067	113	0	-0.078324	SPORT UTILITY'	PASSENGER CAR'	NO	1.50494	-1.313731	0.288377	crash
45	NO SPECIAL FEATU	STRAIGHT,LEVEL'	STATE SECONDAR'	TWO-WAY,DIVIDED,	COARSE ASPHALT'	DRY	DAYLIGHT	CLEAR	NO CONTROL PRES	0.015774	-0.014066	112	0	-0.078324	SPORT UTILITY'	PASSENGER CAR'	NO	-0.128985	-1.313731	-1.476022	injury

Attribute Selection

1) CorrelationAttributeEval
2) InfoGainAttributeEval
3) ReliefF
4) CfsSubsetEval
5) Self-Selected

CorrelationAttributeEval

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Cutoff value of 0.02

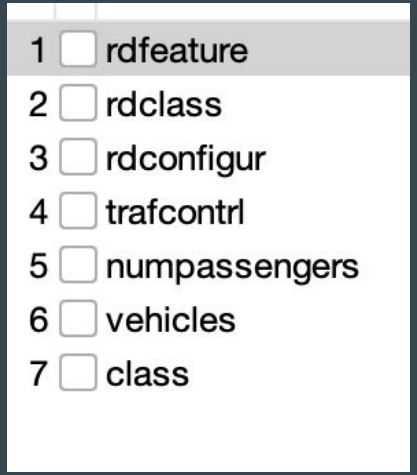
1	☐	rdfeature
2	☐	rdcharacter
3	☐	rdclass
4	☐	rdconfigur
5	☐	rdcondition
6	☐	lightcond
7	☐	trafcontrl
8	☐	zone
9	☐	numpedestrians
10	☐	vehicles
11	☐	class

```
=== Attribute Selection on all input data ===

Search Method:
        Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 21 class):
        Correlation Ranking Filter
Ranked attributes:
 0.128784    9 trafcontrl
 0.11801     3 rdclass
 0.109268    1 rdfeature
 0.107156   20 vehicles
 0.106201    4 rdconfigur
 0.069827   14 numpedestrians
 0.035178   12 zone
 0.029774    2 rdcharacter
 0.025973    7 lightcond
 0.023784    6 rdcondition
 0.014696   16 vehicle2
 0.013492   15 vehicle1
 0.012947    8 weather
 0.012734   13 numpassengers
 0.006762   10 lon2
 0.00541    11 lat2
 0.002679   19 month
 0.00234     5 rdsurface
 0.002267   17 workarea
 0.000728   18 ta_time

Selected attributes: 9,3,1,20,4,14,12,2,7,6,16,15,8,13,10,11,19,5,17,18 : 20
```

InfoGainAttributeEval

Cutoff value of 0.01

$$IG(A) = \text{Entropy}(D) - \text{Entropy}_A(D)$$




```
Ranked attributes:
 0.037212    3 rdclass
 0.023859    1 rdfeature
 0.016529    9 trafcontrl
 0.015225   20 vehicles
 0.014701    4 rdconfigur
 0.011723   13 numpassengers
 0.004443   12 zone
 0.00321    14 numpedestrians
 0.001609   15 vehicle1
 0.001276   18 ta_time
 0.001193    2 rdcharacter
 0.001129   16 vehicle2
 0.001114    5 rdsurface
 0.000905    6 rdcondition
 0.00082    10 lon2
 0.000737   11 lat2
 0.000721    7 lightcond
 0.000557    8 weather
 0.000152   17 workarea
 0          19 month

Selected attributes: 3,1,9,20,4,13,12,14,15,18,2,16,5,6,10,11,7,8,17,19 : 20
```

ReliefF

$$W(A) = W(A) - \frac{1}{m}\sum_{i=1}^{m}[diff(A, \text{nearest\_hit}) - diff(A, \text{nearest\_miss})]$$

Cutoff value of 0.02

1	☐	rdfeature
2	☐	rdcharacter
3	☐	rdclass
4	☐	rdconfigur
5	☐	rdsurface
6	☐	lightcond
7	☐	weather
8	☐	trafcontrl
9	☐	lon2
10	☐	zone
11	☐	vehicle1
12	☐	vehicle2
13	☐	ta_time
14	☐	vehicles
15	☐	class

```
Ranked attributes:
 0.0919194   12 zone
 0.0831862    3 rdclass
 0.0587109   15 vehicle1
 0.0515362   16 vehicle2
 0.0429486    1 rdfeature
 0.0418798    4 rdconfigur
 0.0382455   19 month
 0.0334227    5 rdsurface
 0.0305994    7 lightcond
 0.0287035    8 weather
 0.0221575   18 ta_time
 0.0204764    2 rdcharacter
 0.0203199   20 vehicles
 0.020221     9 trafcontrl
 0.0088835    6 rdcondition
 0.0062019   14 numpedestrians
 0.0015324   17 workarea
 0.0014038   13 numpassengers
 0.0001509   11 lat2
 0.0000133   10 lon2

Selected attributes: 12,3,15,16,1,4,19,5,7,8,18,2,20,9,6,14,17,13,11,10 : 20
```

CfsSubsetEval


```
=== Attribute Selection on all input data ===

Search Method:
        Greedy Stepwise (forwards).
        Start set: no attributes
        Merit of best subset found:    0.039

Attribute Subset Evaluator (supervised, Class (nominal): 21 class):
        CFS Subset Evaluator
        Including locally predictive attributes

Selected attributes: 1,3,4,9,13,14,20 : 7
                     rdfeature
                     rdclass
                     rdconfigur
                     trafcontrl
                     numpassengers
                     numpedestrians
                     vehicles
```

Self-Selected

1) weather

2) trafcontrl

3) lightcond

4) rdcondition (Road Condition)

5) rdfeature

6) vehicle1 and vehicle2

7) numpassengers

8) numpedestrians

Classifier Models

K-fold cross-validation

- Naive Bayes
  - A probabilistic classifier based on Bayes' theorem, assuming feature independence
- J48
  - Builds decision tree by repeatedly splitting data based on attribute with the highest information gain
- RandomForest
  - Constructs multiple decision trees using random subsets of data and features, then aggregates votes
- DecisionTable
  - Summarizes data patterns by building and organizing conditions and corresponding actions in a table
- OneR
  - Selects the single best-performing feature and builds a simple rule based on it

Results (Accuracy)

	Naive Bayes	J48	Random Forest	Decision Table	OneR
Self-Selected	85.12%	86.47%	85.18%	86.44%	86.44%
CorrelationAttributeEval	82.76%	86.47%	84.89%	86.46%	86.43%
InfoGain	85.05%	86.47%	85.74%	86.45%	86.44%
ReliefF	84.44%	86.46%	84.32%	86.43%	86.43%
CfsSubsetEval	85.04%	86.47%	85.76%	86.45%	86.43%

Conclusion and Future Work

- Best model was CfsSubsetEval with Naive Bayes
  - 85.0519% Accuracy
  - 0.85 TPR
  - 0.785 FPR
  - 0.799 Precision
  - 0.850 Recall
  - 0.814 F-Measure
  - 0.114 MCC
- Injury and fatality classes have TP rate
- Tradeoff between majority and minority classes
- Address class imbalance

```
=== Confusion Matrix ===

     a     b     c   <-- classified as
 20304   629    29 |     a = crash
  2920   310     5 |     b = injury
    41     2     1 |     c = fatalities
```

Questions?