Problem Set 02

1. Car Insurance (4p)

Find out (by hand or WEKA) with the simple rule (1R) which attribute best predicts whether a car gets stolen or not.

Sample	Attributes					Stolen
	Type	Location	Price	State	Color	
S01	Sport	Town	High	Good	Red	T
S02	Sport	Country	Medium	Medium	White	T
S03	Sport	Suburb	High	Bad	Blue	T
S04	Compact	Suburb	Low	Good	Green	F
S05	SUV	Country	Medium	Bad	White	F
S06	Compact	Country	Low	Medium	Blue	F
S07	SUV	Town	Low	Bad	Red	T
S08	Compact	Country	Medium	Bad	Green	F
S09	Compact	Town	High	Medium	Red	T
S10	Sport	Country	High	Good	White	T
S11	SUV	Town	Medium	Good	Green	F
S12	SUV	Suburb	Low	Bad	White	F
S13	Compact	Suburb	Low	Good	Green	F
S14	Compact	Town	High	Good	Red	T
S15	SUV	Suburb	Medium	Medium	Blue	F

2. Scotts (6p+1p)

Decide (by hand and WEKA) whether Logan is Scottish based on the following attributes and using a Naïve Bayes classifier. Logan likes shortbread, drinks whiskey and eats porridge but doesn't like lager and doesn't watch soccer. Bonus: use a smoothing technique. (+1p)

Sample	Attributes					Scottish
	Shortbread	Lager	Whiskey	Porridge	Soccer	
S01	No	No	Yes	Yes	Yes	No
S02	Yes	No	Yes	Yes	Yes	No
S03	Yes	Yes	No	No	Yes	No
S04	Yes	Yes	No	No	No	No
S05	No	Yes	No	No	Yes	No
S06	No	No	No	Yes	No	No
S07	Yes	No	Yes	Yes	Yes	Yes
S08	No	Yes	No	No	Yes	Yes
S09	Yes	Yes	Yes	Yes	No	Yes
S10	Yes	Yes	No	Yes	No	Yes
S11	Yes	Yes	No	Yes	Yes	Yes
S12	Yes	No	Yes	Yes	No	Yes
S13	Yes	No	Yes	No	No	Yes

Deadline:

October 17, 2016 at 8:00 AM