Question 1:

a) Express Madame Irma's six statements into First Order Logic (FOL).

$$\exists x\big(dog(x) \land have(You, x)\big) \quad 1$$
$$buyCarrot(Robin) \quad 2$$
$$\forall x \forall y\big(owns(x, y) \land rabbit(y)\big) \to \forall z \forall w\big(rabbit(w) \land chases(z, w)\big) \to hates(x, z) \quad 3$$
$$\forall x \left(dog(x) \to \exists y\big(rabbit(y) \land chases(x, y)\big)\right) \quad 4$$
$$\forall x \left(buyCarrot(x) \to \exists y \left(owns(x, y) \land \big(rabbit(y) \lor grocery(y)\big)\right)\right) \quad 5$$
$$\forall x \forall y \forall z\big(owns(y, z) \land hates(x, z) \to \neg date(x, y)\big) \quad 6$$

b) Translate the obtained expressions to Conjunctive Normal Forms (CNFs).

$$dog(D) \land have(You, D) \quad 1$$
$$buyCarrot(Robin) \quad 2$$
$$\big(owns(x, y) \land rabbit(y)\big) \land \big(\neg rabbit(w) \lor \neg chases(z, w)\big) \lor hates(x, z) \quad 3$$
$$(\neg dog(x) \lor rabbit(R)) \land \big(\neg dog(x) \lor chases(x, R)\big) \quad 4$$
$$\big(\neg buyCarrot(x) \lor own(x, Y)\big) \land \big(\neg buyCarrot(x) \lor rabbit(Y\ ) \lor grocery(Y)\big) \quad 5$$
$$\neg own(y, z) \lor \neg hate(x, z) \lor \ \neg date(x, y) \quad 6$$

c) FOL: $\exists x(grocery(x) \land \neg owns(Robin, x)) \to \neg date(Robin, You)$

Negate it: $\neg(\exists x\big(grocery(x) \land \neg owns(Robin, x)\big) \to \neg date(Robin, You))$

Then convert it to CNF: $grocery(G) \land \neg owns(Robin, G) \land date(Robin, You)$

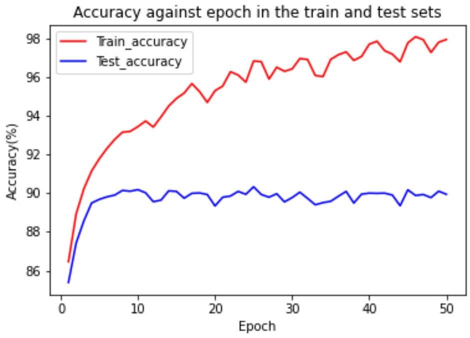
d) Split the negated CNF as $\neg owns(Robin, Grocery)$ and $date(Robin, You)$

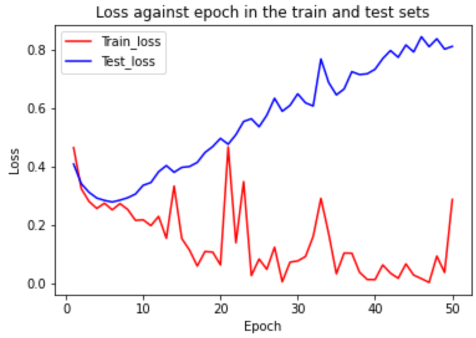
1. Resolve $\neg owns(Robin, Grocery)\ \land CNF5$:
$$\neg buyCarrot(Robin) \lor owns(Robin, Rabbit)$$
2. Resolve the result above and CNF2:
$$owns(Robin, Rabbit)$$
3. Resolve the result above and CNF3:
$$\neg chases(z, Rabbit) \lor hates(Robin, z)$$
4. Resolve the result above and CNF4:
$$hates(Robin, Dog)$$
5. Resolve CNF1 and CNF6:
$$\neg hate(x, Dog) \lor \neg date(x, You)$$
6. Resolve the step 4, step 5 and $date(Robin, You)$:
$$hates(Robin, Dog) \land \big(\neg hate(x, Dog) \lor \neg date(x, You)\big) \land date(Robin, You)$$
$$= \emptyset$$
Proved!

Question 2:

a) Cross entropy loss. It is widely used in classification tasks.

b) The final accuracy obtained on the train and test set is 97.93% and 89.94% respectively. From the first plot below, the accuracy in train set goes up dramatically with the increase of training epoch, whereas that in test set keeps stable after rises at the beginning. From the second plot below, with the rise of the training epoch, the test loss grows up after experiences a slight drop, whereas there is almost a decline throughout the whole period enough though some slight flucuations. In other words, with the epoch increasing, the model performs better and better in the training set, but performs worse in the test set. It can be concluded that there is a overfitting pheonemnon after over training. As the shown in the plot, It may be a good option to only train for about five epoches.





c) Activation function

Final Accuracy	Relu	Tanh	Sigmoid	ELU
Training set	97.93%	99.96%	93.32%	98.86%
Test set	89.94%	91.14%	90.31%	90.25%

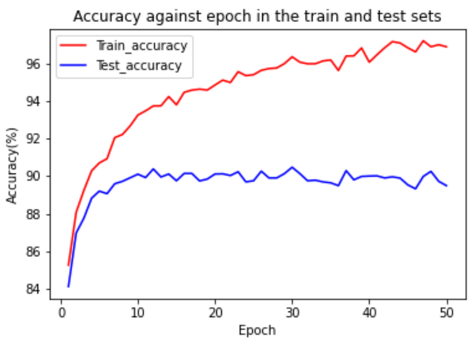
As shown in the table above, it is clear that the performance of the model is the best in terms of accuracy when Tanh activation function is used.

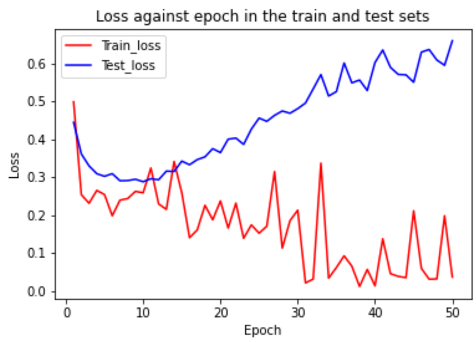
Learning rate

	0.001	0.1	0.5	1	10
Training set	87.12%	97.93%	90.48%	10%	10%
Test set	86.39%	89.94%	85.54%	10%	10%

As shown in the table above, it is clear that the proformance of the model is the best in terms of accuracy when the learning rate is set as 0.1. When the learning rate is too small, the gradient decent can be too slow; when the learning rate is too large, the loss of the model is hard to converge or overshoot the optimum.

d) When adding a dropout of 0.3 on the second fully connected layer, the accuracy of the model in the training data is 96.90% and 89.50% in the test data. Compared to the case without dropout, there is no obvious change on the performance. When increasing the dropout, the accuracy obtained in the training set decrease slightly. It seems that the dropout can decrease the overfitting.





Dropout

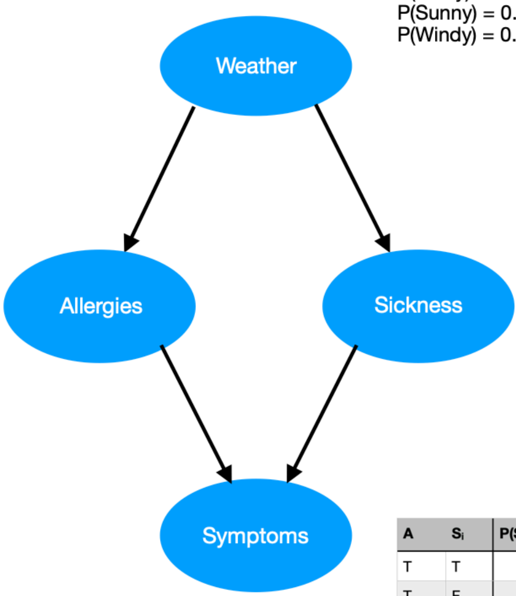
Accuracy	0.1	0.3	0.5	0.7
Training set	97.37%	96.90%	96.08%	95.09%
Test set	90.23%	89.50%	89.90%	89.56%

Question 3:

a) & b)

b) Bayesian Network Add the extra information

P(Rainy) = 0.25
P(Sunny) = 0.6
P(Windy) = 0.15



R	S	W	P(A)
T	F	F	0.15
F	T	F	0.5
F	F	T	0.6

R	S	W	P(S$_i$)
T	F	F	0.8
F	T	F	0.15
F	F	T	0.45

A	S$_i$	P(S$_y$=H)	P(S$_y$=R$_n$)	P(S$_y$=S$_t$)
T	T	0.6	0.3	0.1
T	F	0.75	0.2	0.05
F	T	0.1	0.5	0.4
F	F	0.15	0.05	0.8

c) $P(S_y = H, A = False \mid W = S)$

$$= \frac{P(S_y = H, A = False, W = S)}{P(W = S)}$$

$$= \frac{[P(S_y = H, A = False,\ S_i = True, W = S) + P(S_y = H, A = False,\ S_i = False, W = S)]}{P(W = S)}$$

$$= \frac{P(S_y = H|A = False,\ S_i = True) \cdot P(A = False|W = S) \cdot P(\ S_i = True|W = S) \cdot P(W = S)}{P(W = S)}$$

$$+ \frac{P(S_y = H|A = False,\ S_i = False) \cdot P(A = False|W = S) \cdot P(\ S_i = False|W = S) \cdot P(W = S)}{P(W = S)}$$

$$= P(S_y = H|A = False,\ S_i = True)P(A = False|W = S)P(\ S_i = True|W = S)$$
$$+ P(S_y = H|A = False,\ S_i = False)P(A = False|W = S)P(\ S_i = False|W = S)$$
$$= (0.1 \times (1 - 0.5) \times 0.15) + (0.15 \times (1 - 0.5) \times (1 - 0.15))$$
$$= 0.07125$$

d) $P(S_y = S_t \mid W = R)$

$$= \frac{P(S_y = S_t, W = R)}{P(W = R)}$$

$$= \frac{\begin{gathered}[P(S_y = S_t, A = True,\ S_i = True, W = R) \\ +P(S_y = S_t, A = True,\ S_i = False, W = R) \\ +P(S_y = S_t, A = False,\ S_i = True, W = R) \\ +P(S_y = S_t, A = False,\ S_i = False, W = R)]\end{gathered}}{P(W = R)}$$

$$= \begin{gathered}P(S_y = S_t \mid A = True,\ S_i = True) \cdot P(A = True|W = R) \cdot P(\ S_i = True|W = R) \\ +P(S_y = S_t \mid A = True,\ S_i = False) \cdot P(A = True|W = R) \cdot P(\ S_i = False|W = R) \\ +P(S_y = S_t \mid A = False,\ S_i = True) \cdot P(A = False|W = R) \cdot P(\ S_i = True|W = R) \\ +P(S_y = S_t \mid A = False,\ S_i = False) \cdot P(A = False|W = R) \cdot P(\ S_i = False|W = R)\end{gathered}$$

$$= 0.1 \times 0.15 \times 0.8 + 0.05 \times 0.15 \times (1 - 0.8) + 0.4 \times (1 - 0.15) \times 0.8 + 0.8 \times (1 - 0.15) \times (1 - 0.8)$$
$$= 0.4215$$