Istanbul Technical University- Spring 2017
BLG527E Machine Learning
Homework 4

Omid Abdollahi Aghdam
student number: 504151520
abdollahi15@itu.edu.tr
abdollahi.omid@gmail.com

May 28, 2017

Q1) This is a Head-to-Head connection and we use diagnostic inference (1) to compute the probability. We use the given marginal and conditional probabilities to find $P(G = 0|F = 1)$ and $P(G = 0)$ and the replace them in the 1.

$$P(F = 1|G = 0) = \frac{P(G = 0|F = 1)P(F = 1)}{P(G = 0)} \qquad (1)$$

We calculate $P(G = 0)$ by marginalizing over the joint probabilities.

$$
\begin{aligned}
P(G = 0) &= \Sigma_{B,F} P(G, B, F) \\
&= P(G = 0|B = 1, F = 1)P(B = 1|F = 1) + P(G = 0|B = 0, F = 0)P(B = 0|F = 0) \\
&\quad + P(G = 0|B = 1, F = 0)P(B = 1|F = 0) + P(G = 0|B = 0, F = 1)P(B = 0|F = 1) \\
&= 0.315
\end{aligned}
$$
$$(2)$$

To compute $P(G = 0|F = 1)$, which is a causal (predictive) inference, we marginalize over B.

$$
\begin{aligned}
P(G = 0|F = 1) &= \Sigma_B P(G, B, F) \\
&= P(G = 0|B = 1, F = 1)P(B = 1|F = 1) + P(G = 0|B = 0, F = 1)P(B = 0|F = 1) \\
&= P(G = 0|B = 1, F = 1)P(B = 1) + P(G = 0|B = 0, F = 1)P(B = 0) \\
&= 0.26
\end{aligned}
$$
$$(3)$$

We replace the result of 3 and 2 in 1 to compute $P(F = 1|G = 0)$.

$$
\begin{aligned}
P(F = 1|G = 0) &= \frac{0.26 \times 0.9}{0.315} \\
&= 0.743
\end{aligned}
$$
$$(4)$$

Q2) Observation sequence is $\{O_1 = Red, O_2 = Red\}$, we calculate $P(O|A, B, \pi)$ using forward variable.

Initialization:
We denote state 1 and 2 with S_1, S_2 respectively.

$$
\begin{aligned}
\alpha_{t=1}(S_1) &= \pi_{S_1} b_{S_1}(Red) \\
&= 0.2 \times 0.3 \\
&= 0.06
\end{aligned}
$$
$$(5)$$

$$
\begin{aligned}
\alpha_{t=1}(S_2) &= \pi_{S_2} b_{S_2}(Red) \\
&= 0.8 \times 0.5 \\
&= 0.4
\end{aligned}
$$
$$(6)$$

Recursion:

$$\alpha_{t=2}(S_1) = [\Sigma_{i=1}^2 \alpha_{t=1}(i)a_{i1}]b_{S_1}(Red)$$
$$= [0.06 \times 0.6 + 0.4 \times 0.5]0.3 \tag{7}$$
$$= 0.0708$$

$$\alpha_{t=2}(S_2) = [\Sigma_{i=1}^2 \alpha_{t=1}(i)a_{i2}]b_{S_2}(Red)$$
$$= [0.06 \times 0.4 + 0.4 \times 0.5]0.5 \tag{8}$$
$$= 0.112$$

$$p(O|A, B, \pi) = \Sigma_{i=1}^2 \alpha_T(i)$$
$$= \alpha_{t=2}(S_1) + \alpha_{t=2}(S_2)$$
$$= 0.0708 + 0.112 \tag{9}$$
$$= 0.1828$$

Q3a) We have selected Keras. It is written with python and has Tensorflow or Theano as backend and provides high level API for implementing deep learning models. Using keras we can implement any artificial neural network architecture in a short amount of time.

Q3b) We implement a three layer neural network model using keras (lines 94-101 in file hw4.py). Afterwards, we tune the hyper-parameters (lines 103-127) and train the best model on whole train data. Train, test time, and confusion matrices will be outputted on terminal by running hw4.py. Figures 1, 2, shows the train and test confusion matrices respectively.
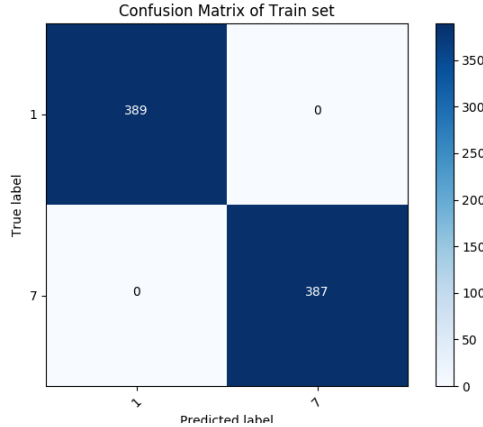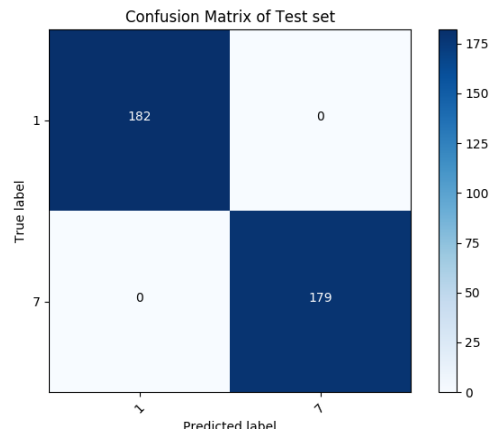

Figure 1: Confusion Matrix of Train set.

Figure 2: Confusion Matrix of Train set.