CSE 535 Midterm Spring 2020 Solution

Q1: Consider a restaurant recommendation application (Live2Eat), which shows you nearby restaurants for a given location. Assume that Live2Eat automatically updates the location as the user moves, and that it also updates the nearby restaurants. It has **two options for obtaining the location information.**

The first option: by using GPS, which is more accurate.

The second option: by using the mobile tower-based cellular network, which is far less accurate.

Suppose that there is a Live2Eat user who is driving down a street, and the GPS signal is lost at time t = 0. Also suppose that the average speed of traffic is 15 kmph. The error in GPS localization is 50 m (0.031 miles), while the error in mobile tower-based localization is 100 m (0.062 miles). Consider that location information is requested by Live2Eat every minute.

When should you switch from GPS to cellular?

Answer:

- The user must switch, because
  - At t = 24s, the user would have covered 100m and the GPS data will be stale. The next time the Live2Eat app requests the data will be at t = 60s.
  - At t = 24s, the localization error for GPS is higher than the mobile tower based localization error.

  (Description with your understanding of the abovementioned key factors are expected. Correct time/ distance calculation of the switch is not significant)

Q2: Imagine a scenario where Boeing designed a pilot monitoring system that deploys in response to a failing MCAS system. If an MCAS system is engaged and the pilot is detected to be stressed, MCAS will automatically disengage.

In this system, consider a brain mobile interface application where the pilot wears a Neurosky headset that senses brain signals (EEG) at **400 Hz. Each brain data point is a 32-bit floating point number. The brain signal is collected by a central controller in the plane and sent to a server, where complex machine learning algorithms are employed to determine the stress level of the pilot.**

Additionally, the aircraft is equipped with sensors, such as the AoA, pitch monitoring, and other relevant sensors. The data rate from the AoA is **5 kbps, and the data rate from the other relevant sensors is **300 kbps**.**

Using the data from these sensors, the MCAS disable system attempts to predict MCAS failures. If the system detects that the pilot is stressed and an MCAS failure is predicted, the auto-disable facility should disable MCAS. The auto-disable feature only has 5 seconds to make a decision after collecting 5 seconds worth of data.**

There are two options for performing all of the related computation: (a) use a GPU server at the control center, or (b) use a fog server that is onboard the aircraft. The GPU server upload speed is **1 Mbps, whereas the fog server upload speed is 5 Mbps. However, the computation speed of the GPU server is **1500 kbps** (in other words, it can finish the computation on 1500 kb of data in 1 second), whereas the fog server has a computational speed of *200 kbps**.
Which option is faster, fog or GPU?**

Now suppose the failure rate of the GPU server is 0.2. This means that 20% of the time the GPU will send a failure message back to the auto-disable system. When this occurs, the system must transfer all of the information to the GPU server again and redo the computation. The time taken to communicate that a failure has occurred is 500 milliseconds. What is the average *total* time taken for both communication and computation to be performed in the GPU server, in milliseconds?

Answer:

- Calculating data sent by all sensors-
    - Neurosky Headset-
      Rate- 400 Hz
      Size – 32 bit
      Information sent in 1 second – 32 * 400 = 12,800 bits = 12.8kb
      Information sent in 5 seconds – 12.8 * 5 kb
    - AoA-
      Information sent in 1 second – 5kb
      Information sent in 5 seconds – 5 * 5 kb
    - Other relevant sensors-
      Information sent in 1 second – 300 kb
      Information sent in 5 seconds – 5 * 300 kb

  Total data to be transmitted - 12.8 * 5 + 5 * 5 + 5 * 300 = 1589 kb

- Data transmission using GPU-
    - Communication time-
      1 mb data is uploaded in 1 second
      1589 kb data is uploaded in – 1.59 seconds
    - Processing time –
      1500 kb is processed in 1 second
      1589 kb is processed in – 1.06 seconds

Total time needed with GPU – 1.59 + 1.06 = 2.65 seconds

- Data transmission using Fog Server-
    - Communication time-
      5 mb data is uploaded in 1 second
      1589 kb data is uploaded in – 0.32 second
    - Processing time –
      200 kb is processed in 1 second
      1589 kb is processed in – 7.95 seconds

Total time needed with Fog Server- 0.32 + 7.95 = 8.27 seconds

So GPU is faster than Fog Server.

- Calculating with GPU failure-

Error rate – 20%
Time taken to communicate that a failure has occurred – 500 ms = 0.5 second

Time taken with GPU considering failure-

0.8 * 2.65 + 0.2 (2.65 + 0.5 +2.65) = 3.28 Seconds

(Considering n times, or similar combination would be considered valid)

Q3: Consider a smart thermostat (NEST by Google for example). The thermostat shuts down the air conditioning system when the users are not home. Based on a predictive model, given the location of the users the thermostat decides to turn on the air conditioning such that the temperature remains nice when the users come in their smart home. The aim is to save energy by shutting down the AC when the users are not home. Assume that there are 2 people in the house. The following hourly location log is available to a mobile application that interfaces with the Nest thermostat:

Person A

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
12 am	Home	Home	Home	Home	Home	Home	Home
1 am	Home	Home	Home	Home	Home	Home	Home
2 am	Home	Home	Home	Home	Home	Home	Home
3 am	Home	Home	Home	Home	Home	Home	Home

4 am	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**
5 am	**Home**	**Home**	**Park**	**Home**	**Home**	**Home**	**Home**
6 am	**Home**	**Park**	**Park**	**Grocery**	**Home**	**Home**	**Home**
7 am	**Home**	**Park**	**Home**	**Home**	**Park**	**Home**	**Home**
8 am	**Office**	**Home**	**Office**	**Office**	**Park**	**Home**	**Home**
9 am	**Office**	**Office**	**Office**	**Office**	**Home**	**Restaurant**	**Park**
10 am	**Office**	**Office**	**Office**	**Office**	**Home**	**Restaurant**	**Park**
11 am	**Office**	**Office**	**Office**	**Office**	**Office**	**Grocery**	**Park**
Noon	**Office**	**Office**	**Office**	**Office**	**Office**	**Grocery**	**Park**
1 pm	**Office**	**Office**	**Office**	**Office**	**Office**	**Home**	**Park**
2 pm	**Office**	**Office**	**Office**	**Office**	**Office**	**Home**	**Park**
3 pm	**Office**	**Office**	**Office**	**Office**	**Office**	**Home**	**Home**
4 pm	**Office**	**Office**	**Office**	**Office**	**Office**	**Home**	**Home**
5 pm	**Office**	**Office**	**Office**	**Office**	**Office**	**Park**	**Home**
6 pm	**Home**	**Office**	**Grocery**	**Restaurant**	**Movie**	**Park**	**Home**
7 pm	**Home**	**Home**	**Grocery**	**Restaurant**	**Movie**	**Home**	**Home**
8 pm	**Home**	**Home**	**Home**	**Restaurant**	**Movie**	**Restaurant**	**Home**
9 pm	**Home**	**Home**	**Home**	**Home**	**Home**	**Restaurant**	**Home**
10 pm	**Home**	**Home**	**Home**	**Home**	**Home**	**Restaurant**	**Home**
11 pm	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**

Person B

	Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**	**Sunday**
12 am	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**
1 am	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**
2 am	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**
3 am	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**
4 am	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**

5 am	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**
6 am	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**
7 am	**Groceries**	**Home**	**Groceries**	**Home**	**Home**	**Home**	**Home**
8 am	**Groceries**	**Home**	**Groceries**	**Home**	**Home**	**Home**	**Home**
9 am	**School**	**School**	**School**	**School**	**School**	**Restaurant**	**Home**
10 am	**School**	**School**	**School**	**School**	**School**	**Restaurant**	**Home**
11 am	**School**	**School**	**School**	**School**	**School**	**Home**	**Home**
Noon	**School**	**School**	**School**	**School**	**School**	**Home**	**Home**
1 pm	**Home**	**Home**	**Home**	**Club**	**Home**	**Home**	**Park**
2 pm	**Home**	**Home**	**Home**	**Club**	**Home**	**Home**	**Park**
3 pm	**Home**	**Home**	**Home**	**Club**	**School**	**Home**	**Home**
4 pm	**School**	**School**	**School**	**School**	**School**	**Home**	**Home**
5 pm	**School**	**School**	**School**	**School**	**Home**	**Park**	**Home**
6 pm	**Home**	**Home**	**Home**	**Home**	**Movie**	**Park**	**Home**
7 pm	**Home**	**Home**	**Home**	**Home**	**Movie**	**Home**	**Home**
8 pm	**Home**	**Home**	**Home**	**Home**	**Movie**	**Restaurant**	**Home**
9 pm	**Home**	**Home**	**Home**	**Home**	**Home**	**Restaurant**	**Home**
10 pm	**Home**	**Home**	**Home**	**Home**	**Home**	**Restaurant**	**Home**
11 pm	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**	**Home**

Given that the location of Person A is office and Person B is School, at time t = 9am and the Nest thermostat shuts down the AC, find the time after which the Nest thermostat should restart the AC. Assume that it takes 15 mins to get the House to a temperature of 75 degrees from any temperature.

Given that the location of Person A and B are not at home, at time t = 9am and the Nest thermostat shuts down the AC, find the time after which the Nest thermostat

should restart the AC. Assume that it takes 15 mins to get the House to a temperature of 75 degrees from any temperature.

Answer:

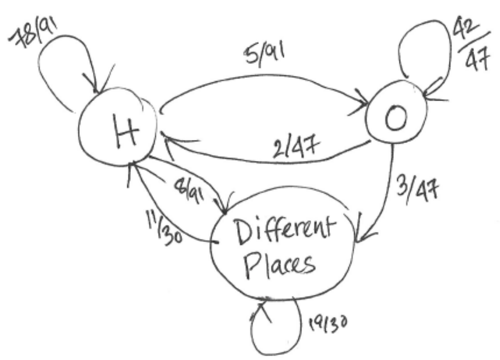
- The idea for this problem is to solve for a stochastic markov model.
- There are several approaches that are correct, provided that the solution is a stochastic one.
- One of the correct approaches is to create markov chains and find stochastic solutions to a steady state.
- Solving the expected time equation discussed in module 2 for Person A and Person B and taking the minimum of those two.

Part 1-

For Person A-

We consider the states Home and Office since they are statistically significant. All different places that are individually statistically insignificant, we consider them together.

Markov Chain



Then we solve the expected time equation for the person returning home, given that the person is at office, $E[T_{H|O}]$

From the above model we get three equations-

$$\boldsymbol{E[T_{H|O}] - 1 = P(O|O) * E[T_{H|O}] + P(H|O) * E[T_{H|H}] + P(D|O) * E[T_{H|D}]}$$
$$\boldsymbol{E[T_{H|H}] - 1 = P(H|H) * E[T_{H|H}] + P(O|H) * E[T_{H|O}] + P(D|H) * E[T_{H|D}]}$$
$$\boldsymbol{E[T_{H|D}] - 1 = P(D|D) * E[T_{H|D}] + P(H|D) * E[T_{H|H}] + P(O|D) * E[T_{H|O}]}$$

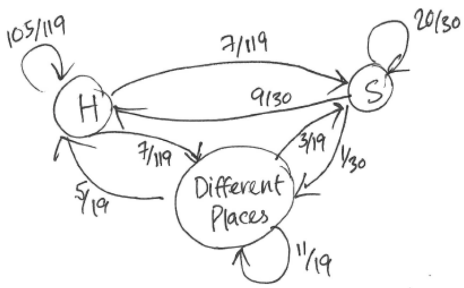
Note that, $\boldsymbol{E[T_{H|H}]}$ **= 0,** and solving the above three equations and using values from the markov chain result in,

$E[T_{H|O}]$ = 11.04

For Person B-

We consider the states Home and School since they are statistically significant. All different places that are individually statistically insignificant, we consider them together.

Markov Chain



Then we solve the expected time equation for the person returning home, given that the person is at School, $E[T_{H|S}]$

From the above model we get three equations-

$$E[T_{H|S}] - 1 = P(S|S) * E[T_{H|S}] + P(H|S) * E[T_{H|H}] + P(D|S) * E[T_{H|D}]$$
$$E[T_{H|H}] - 1 = P(H|H) * E[T_{H|H}] + P(S|H) * E[T_{H|S}] + P(D|H) * E[T_{H|D}]$$
$$E[T_{H|D}] - 1 = P(D|D) * E[T_{H|D}] + P(H|D) * E[T_{H|H}] + P(S|D) * E[T_{H|S}]$$

Note that, **$E[T_{H|H}]$ = 0,** and solving the above three equations and using values from the markov chain result in,

$E[T_{H|S}]$ = 3.36

Now taking the minimum of the two, $E[T_{H|O}]$ = 11.04 and $E[T_{H|S}]$ = 3.36 is 3.36. Taking a next discrete value for Person B to be at home (9am + 3.36 hrs) is 1pm. So, the AC should be turned on at (1:00pm – 15 minutes) = 12:45pm

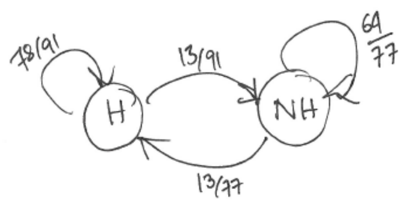
(Taking the discrete value is not necessary)

Part 2-

For Person A-

We consider the states Home and NotHome.

Markov Chain



Then we solve the expected time equation for the person returning home, given that the person is not at home, $E[T_{H|NH}]$

From the above model we get the equation-

$$E[T_{H|NH}] - \mathbf{1} = \boldsymbol{P}(NH|NH) * \boldsymbol{E}[\boldsymbol{T}_{\boldsymbol{H}|NH}] + \boldsymbol{P}(\boldsymbol{H}|NH) * \boldsymbol{E}[\boldsymbol{T}_{\boldsymbol{H}|\boldsymbol{H}}]$$

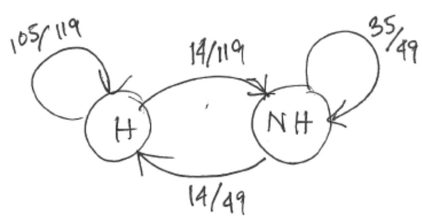
Note that, $\boldsymbol{E}[\boldsymbol{T}_{\boldsymbol{H}|\boldsymbol{H}}]$ **= 0** and using the values from the markov chain results in,

$$E[T_{H|NH}] = 5.92$$

For Person B-

We consider the states Home and NotHome.

Markov Chain



Then we solve the expected time equation for the person returning home, given that the person is not at home, $E[T_{H|NH}]$

From the above model we get the equation-

$$E[T_{H|NH}] - \mathbf{1} = \boldsymbol{P}(NH|NH) * \boldsymbol{E}[\boldsymbol{T}_{\boldsymbol{H}|NH}] + \boldsymbol{P}(\boldsymbol{H}|NH) * \boldsymbol{E}[\boldsymbol{T}_{\boldsymbol{H}|\boldsymbol{H}}]$$

Note that, $\boldsymbol{E}[\boldsymbol{T}_{\boldsymbol{H}|\boldsymbol{H}}]$ **= 0** and using the values from the markov chain results in,

$$E[T_{H|NH}] = 3.5$$

Now taking the minimum of the two, 5.92 and 3.5 is 3.5
Taking a next discrete value for Person B to be at home (9am + 3.5 hrs) is 1pm.
So, the AC should be turned on at (1:00pm – 15 minutes) = 12:45pm

(Taking the discrete value is not necessary)

Q4)
For part A, remember you only have access to world-class and Black box machine that given input produces softmax output. The idea here is to use a strategy that would take a certain number of points from world-class, perform some kind of perturbation to move the point into client class territory. After some number of such perturbation, we will be able to get parameters for wx + b.

For part B, we expect an answer to be NO. For kernel SVM, the above or similar strategies would not work. Solving for kernel function would require a lot more resources and different strategies described in the question. If you answered yes, you would need to provide some more analyses and detailed explanation to be accepted.

Q5)

Given that we have fixed TTL and capacity requirements of 6min and 250 cars, we have two options for extending the range of the tower.

Option 1 is setting the range based on TTL. Given that cars travel 45 mph,

Range = Time * Speed

thus we can travel 6 * (45/60) = 4.5 miles within TTL time. So the range might be 4.5 miles.

Option 2 is to set solve from the standpoint of fixed capacity.

Capacity = Traffic * Range * Time

where Time is a duration we are within tower: Time = Range/ Speed, we assume that cars are removed from the registry if it leaves the range.

Worst-case traffic is 37 per min per mile

Capacity = Traffic * Range * (Range/ Speed) =>

250 = 37 per min per mile * Range * (Range/ 0.75) min, Solving for Range, would give us 2.25 miles.

Given these two options, we would need to get the min(4.5, 2.25) miles. The answer should be 2.25 miles.

There could be some other slightly different solutions depending on a different set of assumptions.