Assignment Nine
ECE 4200

- Provide credit to **any sources** other than the course staff that helped you solve the problems. This includes **all students** you talked to regarding the problems.
- You can look up definitions/basics online (e.g., wikipedia, stack-exchange, etc)
- **The due date is 5/9/2020, 23.59.59 Eastern time**.
- Submission rules are the same as previous assignments.
- **Please write your net-id on top of every page. It helps with grading.**

Problem 1 (30 points). Consider the following movie rating matrix with five users.

	LOTR	HPATPOZ	Snatch	LSATSB	The Gentlemen	The Hobbit
A	5	?	1	2	3	4
B	5	4	2	2	2	5
C	1	2	4	?	4	3
D	?	2	4	5	?	?
E	?	3	5	4	5	1

1. Compute the user-user similarity for all the 10 pair of movies and 12 pairs of users using Pearson's similarity, and Euclidean similarity.
2. Using k-NN user-user CF with $k = 1$, $k = 3$ fill in the missing entries in the matrix above using Pearson similarity.
3. Fill in the missing entries with the answer you obtained in the previous part with $k = 3$. Then using k means clustering with Pearson similarity measure for $k = 2$ obtain a cluster of the movies. You can start with an initial cluster centers as LOTR and HPATPOZ.
4. Using the Euclidean similarity distance, perform a hierarchical clustering using single-linkage clustering to generate a hierarchy of clusters for both the movies and the users.

Problem 2 (30 points). Consider a Markov chain with three states, Overcast, Rain, and Sunny. The transition probabilities are given in the following table. The $(i, j)th$ entry of the matrix is the probability that the next day with be j if today is i. April 29, 2020 is Rain.

	O	S	R
O	1/3	1/3	1/3
S	1/4	1/2	1/4
R	1/4	1/4	1/2

1. Draw the state transition diagram with arrows annotating the transition probabilities.
2. What is the probability that it will be Sunny on April 30th, 2020?
3. What is the probability that it will Rain on May 2nd, 2020?
4. What is the probability that it will Rain every day until May 5, 2020 (including it)?
5. Compute the probability of it Rain on May 6, 2020?

Problem 3 (30 points). See attached Jupyter Notebooks for details.