Assignment 3

1. (5 pts) Classify using Naïve Bayes method on the titanic dataset the data items:

```
2nd     child   male    ?
2nd     adult   female  ?
```

You can use a spreadsheet to compute the counts.

2. (5 pts) Complete the accompanying python notebook on creating a neural using PyTorch.

3. (5 pts) Consider the following set of users and movies they have rated.

`'Lisa Rose': {'Lady in the Water': 2.5, 'Snakes on a Plane': 3.5, 'Just my Luck': 3.0, 'Superman Returns': 3.5, 'You, Me and Dupree': 2.5, 'The Night Listener': 3.0}`	`'Gene Seymour':{'Lady in the Water': 3.0, 'Snakes on a Plane': 3.5, 'Just my Luck': 1.5, 'Superman Returns': 5.0, 'The Night Listener': 3.0, 'You, Me and Dupree': 3.5}`
`'Michael Phillips':{'Lady in the Water': 2.5, 'Snakes on a Plane': 3.0, 'Superman Returns': 3.5, 'The Night Listener': 4.0}`	`'Claudia Puig': {'Snakes on a Plane': 3.5, 'Just my Luck': 3.0, 'The Night Listener': 4.5, 'Superman Returns': 4.0, 'You, Me and Dupree': 2.5}`
`'Mick LaSalle':{'Lady in the Water': 3.0, 'Snakes on a Plane': 4.0, 'Just my Luck': 2.0, 'Superman Returns': 3.0, 'The Night Listener': 3.0, 'You, Me and Dupree': 2.0}`	`'Jack Matthews':{'Lady in the Water': 3.0, 'Snakes on a Plane': 4.0, 'Superman Returns': 5.0, 'The Night Listener': 3.0, 'You, Me and Dupree': 3.5}`
`'Toby': {'Snakes on a Plane': 4.5, 'Superman Returns': 4.0, 'You, Me and Dupree': 1.0}`	

(a) **(3 pts)** Suppose we build a recommender system following the user-user similarities approach with Pearson correlation as a similarity measure. What will be the rating prediction for user Michael Phillips, for movie "You, Me and Dupree"? Give the details of your computation.

In computing the Pearson user-user similarities, restrict the user vectors to only those components (movies) the two users have in common.

(b) **(2 pts)** If we use the user-bias, item-bias approach to recommendation (Netflix competition), what will b_r (short for $b_{\text{lisa rose}}$) be after the first pass over the data? Set $\lambda_1=\lambda_2=\gamma=0.1$, and start with zero bias values.