

ML questions

- how does clustering work, what methods can be used, more about K means, then neural networks, convolutional neural network, what parameters mean(the ones we used in the project) for the convolutional nn, then fully connected neural network, what padding is in fully connected nn, how do you get to the outputs from the nn, what do they represent, gradient descent, how it works, what it calculates, then I got a picture with the LDA and she told me to use Naive Bayes to classify a new point on that plot
- How to use F-1 score as a performance metric if you have multiple classes
How to evaluate how well k-means clustering partitioned the data (I said about the cost function for k-means)
To draw this cost function
To interpret the plot of the data projected onto the first 2 PCs from the project
To interpret the plot of the data projected onto the first 2 LDs from the project
How prediction with Naive Bayes would work for a point with features seen before
How prediction with Naive Bayes would work for a point with features never seen before
- I got the bias variance tradeoff as a topic. They wanted me to explain the formulas behind bias and variance and how we can see that from a plot(which I couldn't fully understand) And then switched to Naive Bayes, how it works, LDA, QDA, which one is more appropriate model and why
- I chose the project, i explained it and they got really into the formulae for LDA reduction and she drew a plot and asked me how would the within scatter matrix look like in the plot or something like that idk and i wasn't sure then the external examiner asked me which other classifiers would we use for our project and i said svm neural networks and from that moment maybe 15 min they asked me stuff about convolutional neural networks specifically
- They asked a lot of stuff about Naive Bayes, CNN and SVM + some stuff about pixel densities

- Picked Logistic regression, got asked questions about Naive Bayes, computing expected loss in a binary setting with a decision boundary at 0.2 instead of 0.5, why is the boundary of logistic regression linear (showing how the log-odds model is a linear function was not enough for them). "How would you make Logistic regression - a discriminative model - into a generative model"
- I got the gradient descent, didn't use notes during the presentation but the moment I wrote the formula on the blackboard, the external guy started to ask me about the notation. I kinda connected it to linear regression too. Then they asked me some details about the types of GD. After that I got the plots of LDA and PCA and how to interpret it. Then what each PC does. This is all I remember.
- Logistic regression
 know the softmax formula by heart
 looked at my A4 notes for the formulas (only during presentation), was fine
 Distribution of y in the likelihood of Logistic regression (binomial, a bunch of either 1s or 0s)
 Explain how to derive the LDA between class scatter matrix
 what does it mean when a convolutional layer is $28 \times 28 \times 32$?
 What does Batch Normalization do? (Only because we had it in our project)
 Notation; in the formulas I wrote, what are we summing over? Had to write subscript i
 For multiclass logistic regression we have a model for each class. Are they trained jointly or separately?
 In the PCA transformed plot, what does it mean for an observation to have a low vs. high value in PC1 and PC2? Relate to the plots of the PCs themselves.