Exam 1

Due No due date **Points** 100 **Questions** 5
Available Oct 9 at 12am - Oct 11 at 11:59pm 3 days **Time Limit** 120 Minutes
Allowed Attempts Unlimited

This quiz was locked Oct 11 at 11:59pm.

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	84 minutes	97 out of 100

Correct answers are hidden.

Score for this attempt: **97** out of 100
Submitted Oct 10 at 5:40pm
This attempt took 84 minutes.

Question 1

20 / 20 pts

Explain how AI, Machine Learning, and Deep Learning are interconnected and how they are different? Be sure and include a definition for each and a timeline. For Machine Learning go into more detail by discussing categories and giving examples.

Your Answer:

AI is computers using logic to make decisions. ML is computers making decisions without being explicitly programmed. DL is computers making decisions without having the decision criteria be explicitly defined. The (rough) timeline is AI (1950s - 90s), ML (1990s - 2010s), DL (2010s to present), but there is significant overlap.

AI, ML, DL are interconnected in that they are require computers and extensive computational power. DL is a subset of ML, and ML is a subset

of AI. They are different in that they require different amounts of human interaction, like in DL humans are more removed from the process.

In ML there are supervised and unsupervised algorithms. Supervised is when there are labels to the data, unsupervised is when the data is unlabeled. Some examples of supervised ML algorithms are linear regression, logistic regression, and decision trees. Examples of unsupervised ML algorithms are SVM (kernel) and KNN.

Question 2

18 / 20 pts

Explain the end to end process for machine learning including why each step is important and what is accomplished in each step.

Your Answer:

The first big step is to get the data. We need to find where to get the data, understand it, and set aside part of it for testing. There are also things we need to be aware of like legal issues of data. we also have to determine the business objective, whether to use supervised or unsupervised model.

The second step is to prepare the data. A important part is handling missing data and often times the median is used. We also want to get rid of things like outliers. Visualizing the data is often very helpful to find patterns and anomalies.

The third step is to train the model. We want to pick a few models that are good and promising. We want to train them and compare their performance based on a performance metric. It's important to find the best model that doesn't overfit or under fit.

The fourth step is testing. We have to test our models on the test set and then tune certain parameters of the models. It's always important to test the model.

The fifth step is to deploy our model and continuously improve. It's important to improve our model as we get more data.

Question 3

19 / 20 pts

Compare and contrast Perceptron, Adaline, and Logistic Regression including what parameters impact the models and what is common between the three models and what is different.

Your Answer:

Perceptron, Adaline, and Logistic Regression are all classification algorithms. Perceptron algorithms works by supplying an initial guess weight vector then, randomly picking a point, and updating the weights, until they don't change. They are different in that the Adaline and Logistic use more of a "soft" classification that (Adaline how much you are wrong, and logistic probability) compared to Perceptron. For Perceptron and Adaline, the weights are parameters that impact the model. For logistic it's the logit function, and parameters are the coefficients

What's common is that all of them are parametric algorithms. Logistic regressions can be extended to more than 2 classes compared to Perceptron and Adaline. Adaline and Logistic both use the gradient descent.

Question 4

20 / 20 pts

Explain the algorithm for each of the following models: KNN, Decision Tree and SVM. What are pros and cons of these types of models? You do not need to answer pros and cons for each model but for this group of models.

Your Answer:

KNN algorithm classifies each point based on the majority category of it's K closest neighbors (based on a distance metric) and the value of K must be picked. Decision Trees split the data based on it's features and uses an algorithm like the ID3 algorithm to pick which feature to split. Often the choice of the feature to split is one that is the greatest reduction in entropy (most information gain). SVM uses kernel methods and the support vectors to separate the different classes, and the decision boundary is the middle of the support vectors.

KNN, decision trees, and SVM are all non-parametric models. The pros for non-parametric models are they are more flexible without having fixed number parameters, can be more powerful, and may perform better than parametric models. Some of the cons of non-parametric models are that they often require more data, tend to be slower, and may be prone to overfitting.

Question 5

20 / 20 pts

Why is Dimensionality Reduction important? What is feature selection versus extraction and what are example techniques for each?

Your Answer:

Dimensionality reduction is important because of the curse of dimensionality. Having a lot of features can lead to worse performance. The number of training examples that are needed also increases greater with more dimensions. Reducing the number of dimensions can be beneficial in that it can reduce the training time, reduce overfitting, and improve accuracy of model.

Feature selection is selecting a subset of the features that filters out redundant, unnecessary features . In contrast feature extraction is creating a brand new feature based on the available information. Techniques like the forward search and backward search (both O(n^2)) are available to find the optimal features for the model. Feature extraction

techniques are LDA and PCA. PCA projects data into a lower dimension and maximizes explained variance. LDA maximizes the separation between the classes.

Quiz Score: **97** out of 100