Worksheet 03

Name:
UID:

Topics

- Intro to DS

Linear Algebra Review

If you need a linear algebra review, please read through the following pdf (https://github.com/gallettilance/CS506-Spring2023/raw/main/worksheets/lecture_03_linear_algebra_review.pdf) before next class

Intro to Data Science

a) what property must a hypothesis have?

There should be examples that falsify the hypothesis.

b) what examples would you have wanted to try?

(4, 8, 16) (0, 0, 0)

c) Poll 1

A

d) Given the hypothesis (x, 2x, 3x), for each of the following, determine whether they are positive or negative examples:

- (2, 4, 6)
- (6, 8 , 10)
- (1, 3, 5)

positive, negative, negative

e) Poll 2

C

f) Describe steps of a Data Science Workflow

- Process Data
- Explore Data
- Extract Features
- Create Model

g) Give a real world example for each of the following data types:

- record
- graph
- image
- text

- record: Bank account e.g. -> (maha, 20, -1000)
- graph: Social networks (social media following/followers)
- image: (Image optimization ?)
- text: Name(text input) -> "Maha"

h) Give a real world example of unsupervised learning

Clustering e.g. Using K-means clustering to classify cat and dog people based on a questionaire, based on user movie ratings give movie recommendations.

i) Give a real world example of supervised learning

- Regression type: Using linear regression model for grade prediction.

Type *Markdown* and LaTeX: α^2