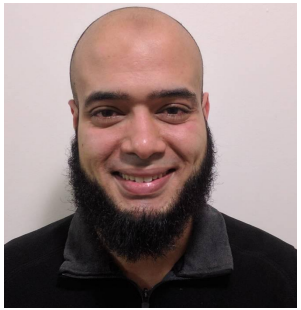
Machine Learning
Coding Interview Questions

Mostafa S. Ibrahim
Teaching, Training and Coaching for more than a decade!

Artificial Intelligence & Computer Vision Researcher
PhD from Simon Fraser University - Canada
Bachelor / MSc from Cairo University - Egypt
Ex-(Software Engineer / ICPC World Finalist)



EDA Questions

- Read data from pandas (e.g. sales.csv) and plot something (e.g. sales trends over time).
  - Think what is a good plot and justify
- Create plots to get insights for a dataset containing the prices of houses and various features like the number of bedrooms, square footage, location, etc
  - **Scatter** plots for numerical features (e.g. price vs square footage)
  - **Bar / box** plots for categorical features
  - For outliers, I would consider using techniques XXX
  - For missing data, I would either impute the missing values using XXX

Write ML Code for

- KNN
- K-Means
- Linear / Logistic Regression
- A 3-layers neural network for regression/classification
  - Add dropout layer (DNN)
- Simple DNN for image classification / Autoencoder for image reconstruction
- Write a **softmax** function (handle the overflow and justify)

Computer Vision Coding Questions

- Write a code that computes intersection over union (IOU)
- Given a 2D image, do flood fill to find CC
- Given a 2D image, rotate the image by **90 degrees in-place**
- Given a 2D image, implement sub-matrix summation for Q queries
    - You will be asked for O(1) per query

Probability Questions

- Random Sampling
  - LeetCode Random Pick Index
  - LeetCode Random Pick with Weight
  - Given a stream of numbers, sample 1 numbers uniformly
    - Harder version: sample k numbers uniformly
    - This is called Reservoir Sampling: video proof code
- Given the following table of XX, compute
  - Compute P(A), Conditional P(A|B), Baye's Probability P(B|A)

Streaming Questions

- Reservoir Sampling
- Given stream of vectors each of D features, compute the **mean and covariance** matrix for each streamed vector
  - First, Compute them if you have the full data
  - Next, try to develop the streamed features.

From Internet

- Implement Matrix Multiplication
- Naive Bias Model

"Acquire knowledge and impart it to the people."

"Seek knowledge from the Cradle to the Grave."