**22-year old coaching client landed a new-grad Machine Learning job at a NYC startup that pays $175k per year ($120k base, $40k stock, $15k annual bonus)**

**Here's the 17 interview questions they were asked: (how many of these could you answer?)**

## Machine Learning
1. What is PCA, why is it helpful, and how does it work?
2. What is heteroskedasticity, and how can you check for it?
3. After loading the Boston house prices dataset (it comes with scikit-learn), can you find any outliers in the data? What should we do with these outliers & why?
4. On the Boston house prices dataset, can you build a simple regression model, and interpret the model's performance?


## ML System Design
5. Imagine you're building a system to recommend users items similar to the items they've bought before. How would you go about building this?
6. For that recommender, how would you handle a new user who hasn't made any past purchases?
7. Let's say you had to scale this recommender to 1 million users, the product catalog had 100k items, & 99% of users had bought less than 3 items before. How would this change your answer?
8. What are feature and concept drift? Give me an example related to the product recommendation system earlier. How would prevent this drift issue?


9. Given a list of integers called nums, return all the triplets [nums[i], nums[j], nums[k]] such that i ! = j, i != k, and j != k, and nums[i] + nums[j] + nums[k] == 0. Example Input: nums = [-1,0,1,2,-1,-4] Example Output: [[-1,-1,2],[-1,0,1]]

10. Given a string s, find the length of the longest substring without repeating characters. Input: s = "abcabcbb" Output: 3 (the answer is 3 because the longest string is "abc")

11. Given a Pandas data-frame of integers, with m rows and n columns, if an element in the data-frame is 0, set its entire row and column to 0's.


## Miscellaneous
12. I see you listed Kubernettes on your resume. What did you use it for & how was your experience? What frustrated you?
13. I see you listed TensorFlow on your resume. What do you like about it? What would you change about TensorFlow?

14. Walk us through your past internship project? What was the most challenging part of getting your model deployed? 15. Walk us through your undergrad research project. Why did you pick XGBoost? What other models did you try?

16. I saw you took Linear Algebra in college. What is an eigen-value? What about an eigen-vector? Can you give me an example of an ML technique that makes use of these concepts?
17. Why do you want to work at this company? Why a startup? Why not grad school?