Suhail Bhat – Algorithmic Interview

**Interviewing the head of the consultancy:**

-- I would start by asking how long did it take you to finish this project? Is this the first such predictive system your team designed?

-- What kind of data was used to train this system?

-- Who collected the data and how big was the dataset? Did you use the Journalism School data or did you use any experimental or observational data?

-- Can you describe the dataset a little more? What were the different columns and rows?

-- How did you translate writing samples, resume and other variables into measurable data?

-- Were there any missing values in the data? If there were, can you tell me how you made sure that this doesn’t result in an error in the system?

-- How does the system weigh different variables such as writing samples, writing test results, demographic data, resumes, and academic transcripts? I mean, do these variables get weighed equally or differently? Also, did you decide to assign weight to these variables or was it the School that gave you its priorities?

-- How does the system compare different variables? How does it know whether a writing sample is good or bad? I am asking this because different people have different writing styles and use different action verbs. How do you measure that? How do you standardize that? Same goes for resumes. Can you please explain this for each of these variables?

-- How much weight have you given to demographic data in this model? My understanding is that 15 or 20 years back, the J School wasn’t as diverse as it is today. This means that the most successful students of the school would have been white and male. How much bias does it create in the system? Did you take measures to rectify this? If yes, can you walk us through how you made sure the system does not carry any racial, ethnic or gender bias?

-- How did you test this system? What is its prediction accuracy?

-- Have the system been updated since it was completed?

-- Is there a chance that the system can select a bad application and fail a good one? If so, what can be done to rectify that? If not, how often can that happen?

**Interviewing the head of admissions:**

-- What are the different variables the School considers in an application? What are the important ones and not so important ones?

-- How did the school decide who is the most successful candidate? What do you mean by successful? What different parameters are used to judge an application?

-- Also, how far back does the list of successful students used in this model go?

-- You have been the head of the admissions for years now. In your experience, how close is this model to the conventional admission process?

-- The variables considered by this system means it is likely going to pass students with great schooling. Would it be fair to use the same barometer for both poor and rich applicants? Doesn’t that create a bias towards poor kids who might not have great grades and writing skills in comparison with rich kids, but have shown tremendous progress in their lives?

-- What happens if the system fails a good student because it couldn’t understand a writing sample due to its unique style and passes a bad one?

-- How important is the demographic data in the admission process? What guidelines were given to the consultancy about that?

-- How close were the predictions when the system was tested? During testing, did you find anything you wish could have been better?