(MSRF) MICROSOFT RESEARCH FELLOWSHIP

—

Interviews

I received a mail for shortlisting on April 5, 2023, from Ravishankar Krishnaswamy. Wherein he discussed the proposed problem statement [Filtered Disk-ANN] and scheduled the interviews for April 24–25. Apart from him, the interview was taken with a panel of four researchers [Harsha, Gopal, and Suresh] individually. Latter also proposed a separate problem statement and was asked to decide among the two.

Ravi	10-11 am	24th April
Gopal	11-12 noon	24th April
Harsha	8:30-9:30 am	25th April
Suresh	2-3 pm	25th April

24th April 2023 - Ravi

The interview started at 10 a.m. on the online Microsoft platform. The interview started with an introduction of the interviewee. [Go for a technical interview rather than a holistic one; they are researchers looking forward to a technical background only.] He then started to ask me about my publications and briefly describe them. I mentioned image segmentation and supervised learning (SGLR). [There was not much clarity about what the project was about.] [Always start with the topic of the project: what was the problem statement, and what was your approach to solving it?] He empathised with me and described the latter project in depth. As I started, he focused on me, not to use technical terms but to address him as a layman and to clarify each and every step of the problem faced and solution derived. He first stardata waswite input data.

What kind of data it was, size, RGB or black and white, which model were the results on, how many layers and number of neurons in each layer, and what is the division between the server and client sides? Specifically, he also asked what is resnet and how you define loss, computation of the cross-entropy loss function and formula, activation function, why argmax in softmax, how weights are being updated in normal resnet architecture, forward and backward propagation, weight update why and how, and what does learning rate signify. Then how does your method work? What problem does it solve? What were the key ideas? What was the metric or application of such a use case? Legit everything from the smallest to the largest detail!

He then summarised how others are going to be focused on coding and algorithms subsequently, along with Suresh giving another project. He then described his problem statement and methodology, to which I synchronized by telling about my information retrieval course and knowledge. He said he had to interview for the rest of this week and then provide the results.

24th April 2023 - Gopal

He was focused on the coding section and asked me about my familiarity with c/c++ and Python. He then asked me to code the strcpy() function without the use of the string library and compile it as I screen-shared. Then she asked about my approach and method and looked at how I came up with classes. He then asked questions on classes, variables, and methods.

Later, he asked about stack trace and stack frame, to which he told me how they were similar to the runtime of Python and gave me a set of bugs and their associated functions. My task was to group similar bugs. I started by discussing k-means and explaining how distance can be seen as a similar function (iou), and then how weights can be added. He then made me code for iou of 2 bug sets, starting with an order of on 2 with 2 for loops, then looking at the sorting o(nlogn) approach with 2 pointers, and finally, the mapping approach, analyze each approach and code the iou metric. Do a dry run on the schema and analyze for errors in the calculation in the case of out-of-bound inputs. We then discussed the day-to-day activities and ended the conversation.

25th April 2023 - Harsha

He was a foreign-based researcher inclined to focus on algorithms and the basics. He started by asking for my CV and asking about the CS courses I took. I told you about AI, ML, DL, information retrieval, NLP, CAOS, algorithms, and data structure. Which he told me not to take AI courses into consideration, only core CS ones. Of which, he told me that for the project, we would need IR, CAOS, linear algebra, and algo. He started by asking about my understanding of the project and the problem statement discussed so far. He told me that what I knew was very

small and that the problem was bigger. He had been the primary person involved in the project with Ravi and thus knew it in detail. He asked me any questions I had about how the work would be executed.

He stated that a good RF would complete two projects, while an average RF would complete at least one. He told me that the primary thing here would be to put aside Python ML things and look into code quality and speed. Get results and publish a paper. attend the conference and get code published on git main merging for the application. There would be ample data and resources available.

Then he asked about his detailed understanding of Diskann's work. Which I described as the motivation for increasing recall and reducing the search time. Which method of storing information in SSD rather than main memory is faster for retrieval, update insertion, and deletion in indices? To which he replied that he needed exact answers and was not going around the bush as an overview. I started by explaining that in SSD memory, the person would organise documents into groups (buckets), to which he questioned that as he worked on it, he was confused about how this came into being. I said that it was the GPs of documents, so when data comes in, the centroid is first evaluated, then the GP is checked. He, like you, might be explaining other algorithms; tell any algorithm in detail, every detail. I talked about hierarchical HNSW, where the source representation of the graph is to be reached to the destination in a greedy way and traversed through layers. To which he asked how to know which greed was good. I was like, random search nearest and then go in the most similar, which is the metric, maybe similarity cosine types. He was unsatisfied, and there was a silence showing the same.

He then asked about computer architecture, about malloc, memset, and memcpy, and how 1 MB of data can be copied, how much time that will take, and to calculate that time, what all values are needed. Which I didn't know. He then asked aboutwhat things I learned in linear algebra that I forgot because it was two years ago. I answered so-ething in differentialuntil I reachedations and laplce transform, which shocked him, and he was like, Okay, tell me about laplace transform. I said from the time domain to the s domain of frequency analysis. He said tell the formula wherein I forgot the s term, so I told $f(s)=integral(f(t)exp(-s*t)dt)$ forgetting the s here. Which frustrated him, and he kept his hand over his head. Then he went to algorithms, and there he asked what was taught. I said sorting, greedy, and graph-based algorithms. He asked about DFS. I did a depth first search, looking till the root node. He asked what data structures were used here. I told her to stack as first in, last out. He said there was another data structure too, but I don't know. He told me how to know when to stop. I said to add a threshold of depth or do early stopping with visiting arrays.

He then asked what part of the project I was interested in coding algorithms for and what was interesting to me. I told him it was the coding part, and he said no ML here, only from scratch algorithms. No learning and modelling here. And the interview was over. One of my worst interviews, with not much knowledge of the domain and being restricted to not getting into the domain (ik ml dl), complete silence within questions, and frustration on his face that I still

remember, was completed like within 35 minutes. He was straightforward and unwilling to entertain any brief answers.

25th April 2023 - Suresh

It was a complete coding round. he was looking for me in his applied ml role. I wish this was done earlier and as primary project selection. It had like 2 programming questions in 1hr where I explain the initial idea then work out the optimization and code the whole thing. First was to find number of i,j pairs where "and" operation gives larger value than "xor" operation of numbers. I came from o(n2) to o(n) complexity. Then there came question on finding the min column value of row sorted binary matrix where i went from o(n*m) to o(m*logn) however it could be reduced further which I couldn't tackle. He then asked me if I had any questions and wrapped up the meet.