Preparing for Google cloud professional machine learning engineer, path for data engineering to be Data Scientist.

# About Myself

My Name is Namit Sehgal, 18+ Years of hands-on experience as extremely passionate DevOps SRE / Database Solution Architect / Data Engineering in both, RDBMs and NoSQL database world. Responsible for full databases life cycle starting form finding optimum database specific appliances, performing server setups using automation (Terraform), database designs, implementations of various project with focus on optimization, protection, database integrity and disaster recovery, releases using CI/CD pipeline tools. Trained and navigated other DBAs, Data Engineers. A solution Architect in Azure and GCP. Setting up data pipelines in cloud providers for transformations.

And.

Changed my role from DBA to Application DBA 10 years back.

Application DBA to Cloud Solution Architect, SRE, Data Engineer 2 years back.

And

I started my cloud journey 5 years back and now doing project management, Solution Architect, Devops, SRE, APPs integration on Cloud and now stabilizing my career or looking ahead myself to be a Data Scientist after being expert Cloud Data Engineer for couple of years.

And still want to have Data as core component in my career.

And

I love to train my Daughter in Maths and she is really doing in many of the Maths olympiad exams and my son who is two years old is becoming naughtier day by day.

My Daughter who is 9years old now, has also started to learn Python in Year Dec 2021.

And

I love eating different food, favourite is Thai.

And

My wife may not agree but I was the one who helped her in the cooking, now making breakfast over the weekends and started my journey in Data Science six months back with my son screaming at my back. So still was able to concentrate and achieve my goal.

No helper for last one year so needs to manage my kid and work working from home but it is fun.

And doing Yoga, running to stay healthy.

# Journey for a Data Engineer in Cloud to be Data Scientist

With the evolution of cloud, I believe Data Engineering is going to be the skills all Data Savvy people should acquire. Their main role is to bring data to Data Scientist so Data Scientist can build their AI models on the quality data and contribute the revenue for Business. This can not only be restricted to Business revenue but can be a free service like health, transport, weather forecast where people do not pay but stilt hey get the output they want.

But you know creating or coding a model just is a 5 percent exercise in full data life cycle. 95 percent of the work is done by Data Engineers.

I am not going to mention here various phases in a data pipeline to produce a working continuous training model as you will have that covered in your trainings.

# Why GCP machine learning engineer?

I thought of Data scientist is because every year I go to Singapore National eye hospital for my retina health and they take a lot of pictures of my eyes to see how different is my retina compared to last year to see how much degradation it has gone through. Even though they could not find much looking with human eyes, but they can use an AI model to do comparison. I thought to compare all such set of images. And tried to explorer on that, there are existing eye macular degradation models available and Data Science is doing good in health sector.

Second reason is being a part of Engineering, I was asked to help my IT service providers with anomaly detection and root cause analysis for operations to make their job easy. I do not know how much more easy job they need though. That included both infrastructure and Application logs for some critical applications to be looked at. So, AI model can help detecting the anomaly and also can help to find out what caused it.

I am working extensively in Azure and but no certifications. I will do those for sure and my MS colleagues keep encouraging and asking me to do. I am sure I can clear those without reading any courses. I will do my DP 203 and DP 100 this year for sure.

Google was the one who wrote paper on Big data, they are the ones came up with Big Table and later with Big Query. Intelligent people took their paper and created other NoSQL databases. So why not start my journey with the creator of handling problems of Big Data.

Google started TensorFlow and later they made in open source. Google made the most intelligent AI systems, you can think of YouTube, Gmail, Google Search, Traffic estimation, many health systems. You will get to know when you do the GCP training on ML.

I did not have any experience to do this exam, Google say to have minimum 3 years exp. In Data Science. But I would say if you are dedicated and motivated, you can do it.

# Steps

* You might have seen this: -

[Professional ML Engineer Exam Guide | Certification  |  Google Cloud](https://cloud.google.com/certification/guides/machine-learning-engineer)

* Note that from Feb 2022, they are going to change AI Platform to Vertex AI in Exams.
* With no experience in ML and AI, I was worried if I can do this but I would say it took me almost 4 months, watching more than 50 hours of training and doing more than 50 labs at [Qwiklabs](https://www.qwiklabs.com/).

<https://googlepluralsight.qwiklabs.com/profile/activity?page=2>

* Following courses at Pluralsight or Coursera:-

Preparing for Google Cloud Certification: Machine Learning Engineer

Advanced Machine Learning on Google Cloud

* Do practice here and look for my comments.

[Free Exam Prep By IT Professionals | ExamTopics](Free%20Exam%20Prep%20By%20IT%20Professionals%20|%20ExamTopics) <https://www.udemy.com/>

* Total number of question 60 and should be done in 2 hours. I was a bit scared but I could do it before time and also review my answers.
* Exam will have 20 to 30 percent question on Google AI services like AI Platform, Kubeflow, BQML, Dataflow pipelines.
* Exam will have 60- 70 questions on scenarios based on general Data Science concepts for example, what will you do improve performance, optimization, models to select, feature engineering, hyperparameters.
* Exam will have 10 percent question on Terraform codes. So be prepare with TensorFlow, python, Pandas or maybe Apache Beam. I would say if you are Data Engineers and are working on Databricks, codes should not be a problem for you.
* After doing so many coding languages, I think now I can learn any language. My society chairman asked to learn Thai with him.
* You must make up your mind that questions can be general fundamental questions, not only related to Google services. For example, you can see in Exam: -

1. Selling a house, price goes higher with latitude and longitude of the house, so study features Engineering.
2. You are using Google Home and you give an instruction to switch off your fan, which Model in GCP is the best? All answers seem correct…Choose the best suitable one. Don’t worry about result.

* In all your practice exams in above courses, make sure you score 60 percent in first attempt. Answers which are wrong, understand Why?
* Next attempt target to have above 90 percent, but do not do immediately as you will remember the answers, switch to different course once you attempt first exam.=
* Please find attached my Notes, I am sure you can add your own notes once you do the course but read it before 1 hour you go for Exam.
* During the Exam, giving outside, you will be asked to wear gloves and cannot remove Mask. Do not panic.
* Mark all answers which you have any doubts for review. In my case, I marked 46 for review as I was sure for only 15 answers.
* No idea how much is passed percentage, so target to score 80 percent in your practice exams. My internal Google resources say > 80 percent is passing score.
* 2 hours is not enough, so end of the Exam, give Google feedback.
* Use bigger font size; you can adjust in the beginning of exam. Bigger make it faster to read, better for my eyes.
* Do not rush for Exam and take atleast 3 months for preparation, It is not AWS or Azure exam which you can do in 10 days preparation and many people are going for these as they are easier to crack. Try something which is challenging and which is going to provide you data solutions for various problems.
* Learning Al and ML fundamentals is not easy. Watch YouTube for things like evolution of neural networks.

[StatQuest with Josh Starmer - YouTube](https://www.youtube.com/c/joshstarmer)

* Watch videos at youtube.com for Google Cloud services for example Big Query, Spanner etc.