



Yogesh Haribhau Kulkarni

Principal Architect (CTO Office, Icertis) | PhD in Geometric Modeling | Google Developer Expert (Machine Learning) | Rasa Hero (Conversational AI) | Neo4j Ninja (Graph Data Science) | Mentor (Mid-career Transition)

[View full profile](#)



Yogesh Haribhau Kulkarni • You

Principal Architect (CTO Office, Icertis) | PhD in Geometric Modeling | Google Devel...
7mo •

Sparse Experts, still better?

Wouldn't it be better if 100s of experts are available to solve your problem than just one?

Even within these 100s of experts, wouldn't it be better to have a way to select just a few, based on their expertise?

Thus, here you get width as well as depth of expertise.

But then what's this 'Sparse' thing?

Sparse Expert Model, dynamically selects experts (constituent-models) based on the input and combines their results into the final one.

For more details, have a look at:

- Tutorial + Interview: <https://lnkd.in/dascNg9h>
- Switch Transformers: Scaling to Trillion Parameter Models with Simple and Efficient Sparsity (https://lnkd.in/d94sbr_X)
- GLaM: Efficient Scaling of Language Models with Mixture-of-Experts (<https://lnkd.in/d8bSZ6sd>)
- Designing Effective Sparse Expert Models (<https://lnkd.in/d5hkXSDv>)

[Yannic Kilcher](#), [Barret Zoph](#), [Irwan Bello](#), [Sameer Kumar](#), [Nan Du](#), [Yanping Huang](#), [Jeff Dean](#), [Noam Shazeer](#) [William \(Liam\) Fedus](#)

[Google Google Developers Experts](#)

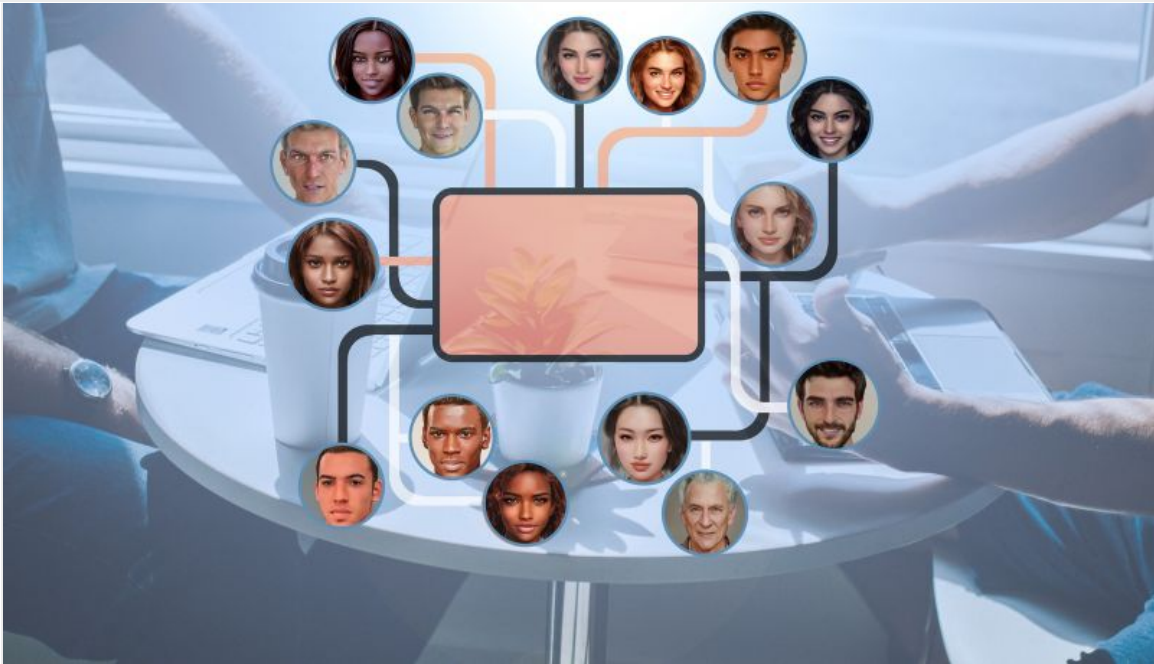
[#SparseExpertModel](#) [#machinelearning](#) [#deeplearning](#) [#artificialintelligence](#) [#ai](#)
[#neuralnetworks](#) [#nlp](#) [#naturallanguageprocessing](#) [#language](#) [#models](#) [#ml](#) [#google](#)



Home

My Network

Jobs



Irwan Bello and 10 others

Reactions



Like

Comment

Repost

Send

3,691 impressions

[View analytics](#)

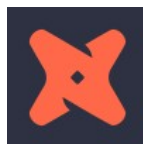


Add a comment...



Ad ...

The modern way to build analytics datasets



Learn the basics of dbt in 30 mins

[Learn for Free](#)



Home



My Network



Jobs