# 5 Free Platforms to Collaborate on Machine Learning Projects

A list of platforms that allows you to collaborate on source code, data, models, metadata, experiment results, pipelines, and more.

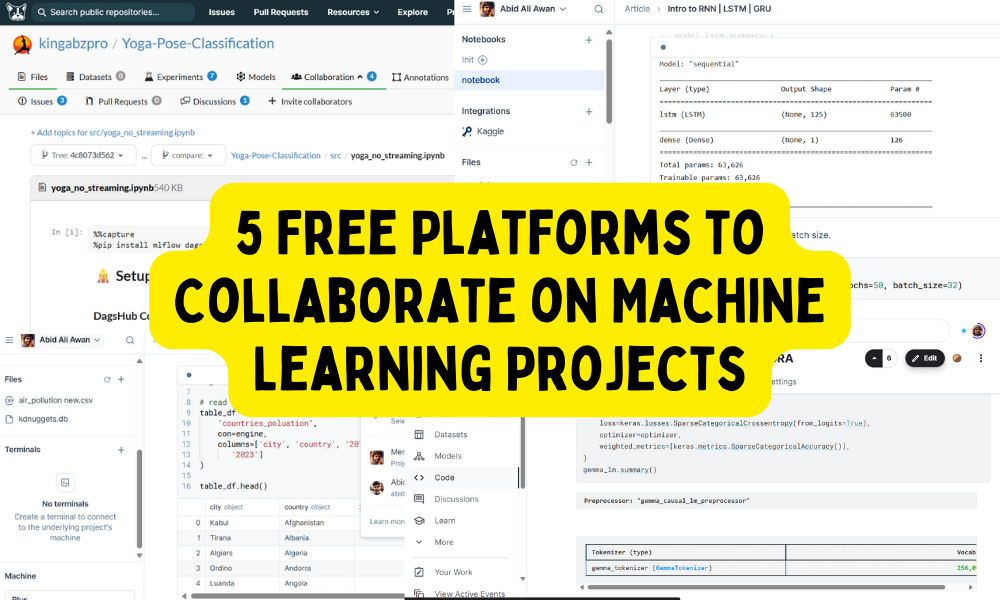


Image by Author

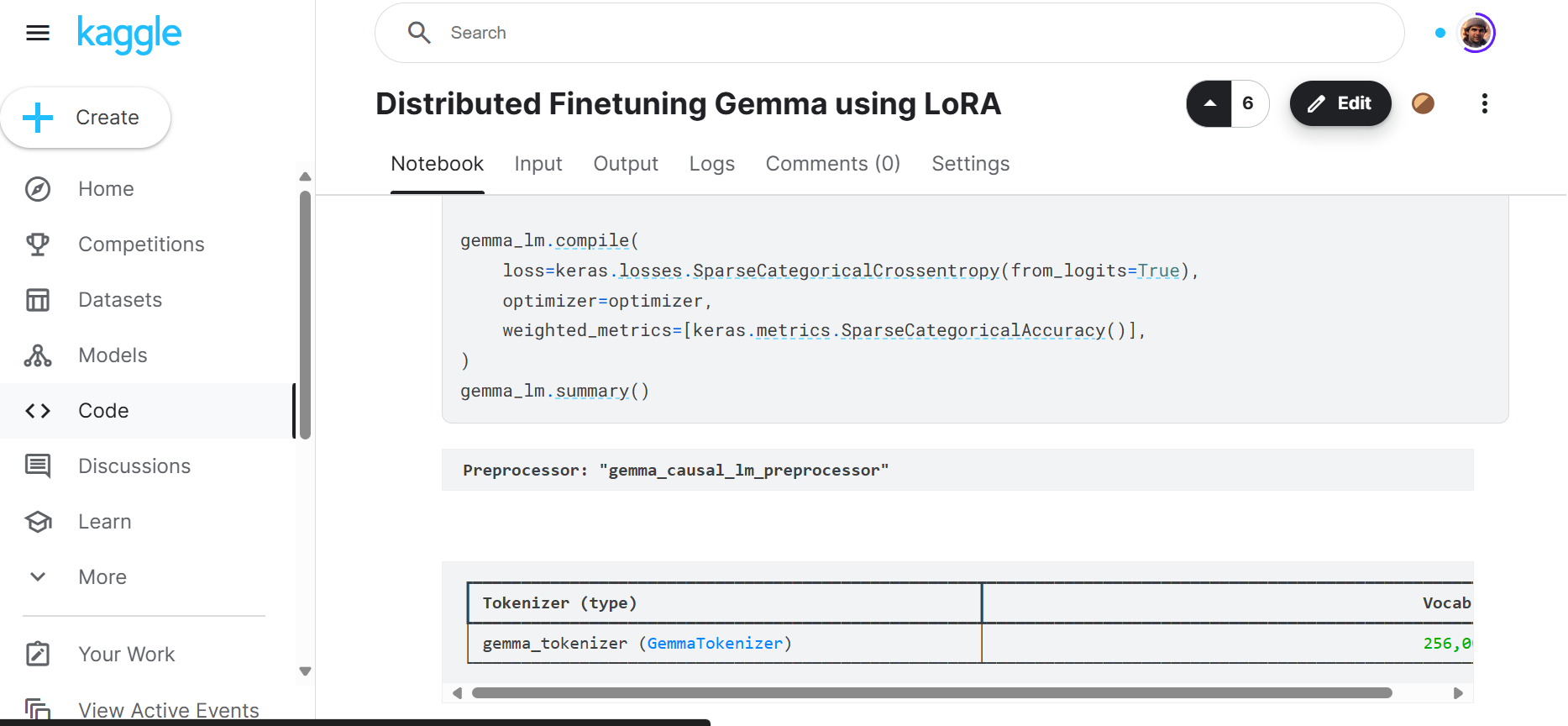
Collaborating on a machine learning project is a bit different from collaborating on a traditional software project. In a machine learning project, engineers are working with data, models, and source code. Additionally, they are also sharing features, model experiment results, and pipelines. You can't just use any code-sharing platform for a machine learning project; you need to select a platform that lets you share code, resolve issues seamlessly, maintain lineage, handle data and models, provide experiment tracking features, support workflow automation, offer external integrations, and more.

As a machine learning engineer, I am always on the lookout for better tools to improve my workflow. In this blog post, I will introduce you to five free platforms designed specifically for collaborating on machine learning projects.

## 1. Kaggle

Kaggle is a well-known platform for machine learning collaboration. It provides access to a vast array of datasets, models, and source code, and also allows you to participate in community discussions and competitions.

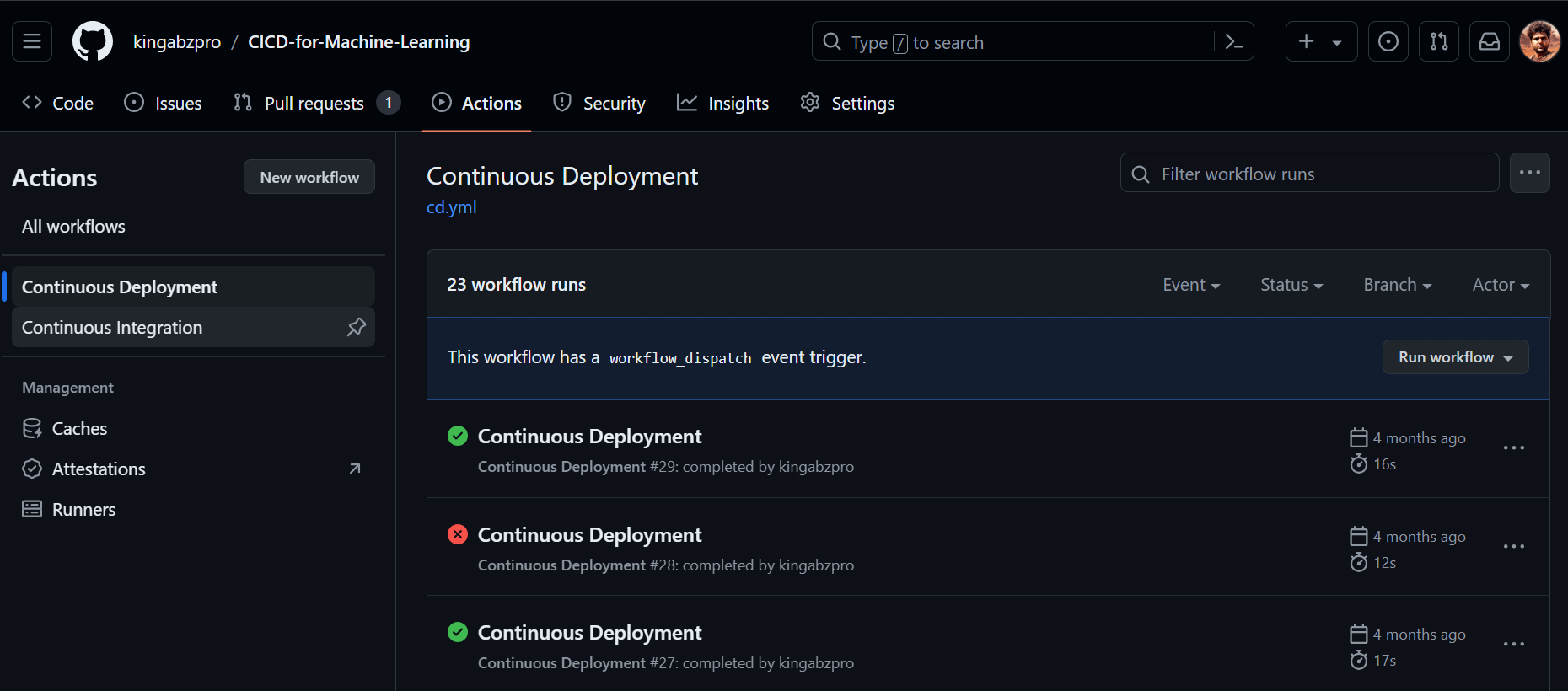
You can keep your data, models, and notebooks private, share them with specific individuals, or make them public. The Kaggle notebooks allow for real-time collaboration on the same notebook. Each notebook includes a comments section, logs, output files, and links to input datasets or models.



The best part of this platform is that it caters to both professionals and beginners, and it is absolutely free. You also get free CPU, GPU, and TPU compute every week, enabling you to train your machine learning models using datasets from Kaggle and save the trained models.

## 2. GitHub

GitHub is the most popular code sharing and collaboration platform, particularly among machine learning engineers. It allows users to share code, models, and large metadata with their team, powered by Git, a decentralized version control system. GitHub also facilitates continuous integration and continuous deployment (CI/CD) using GitHub Actions, enabling engineers to automate their model training, evaluation, and deployment processes.

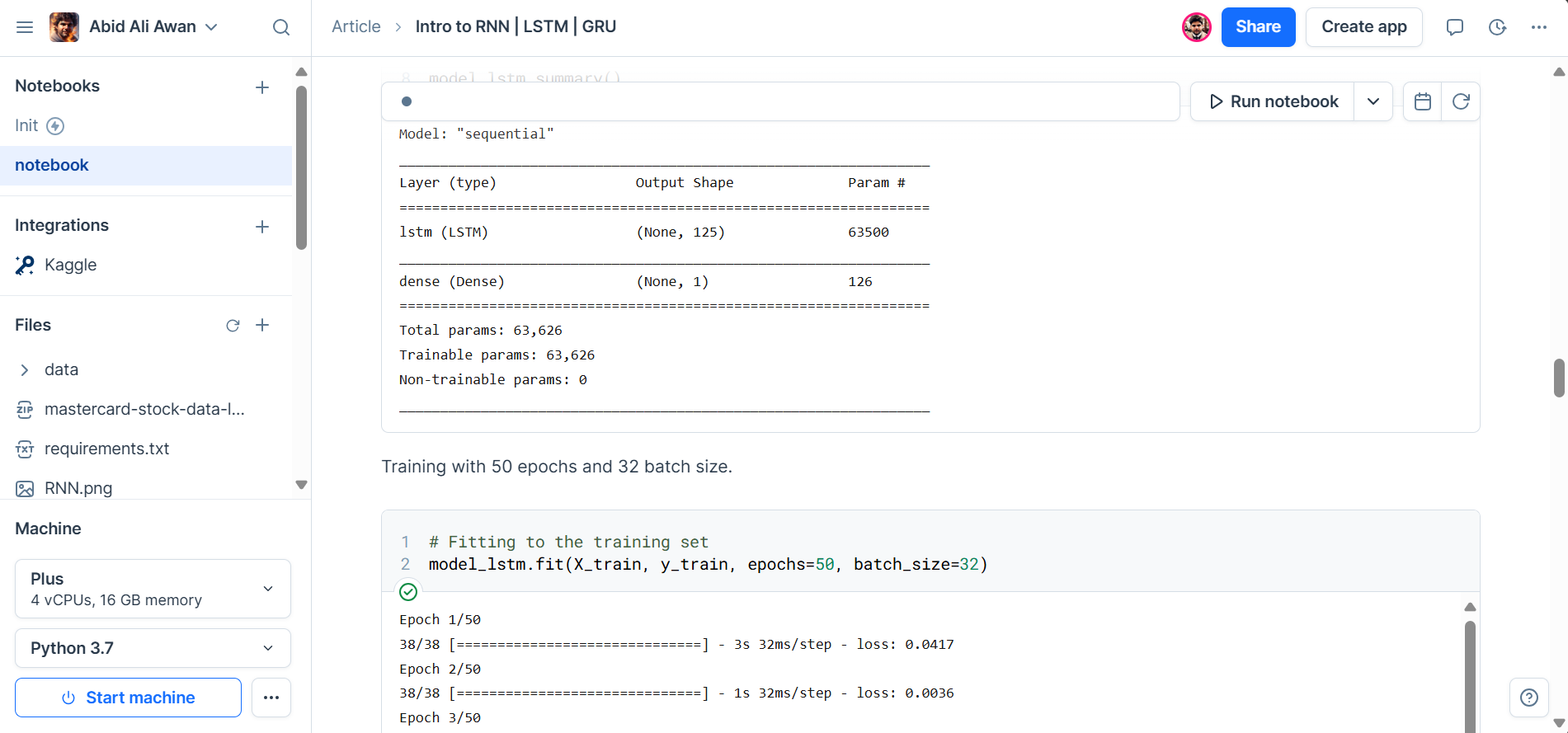


In addition to code sharing and collaboration, GitHub provides features to track project progress, gain insights, and resolve issues faster. The platform also offers CodeSpace, a cloud-based development environment that allows you to make edits and run the code directly from the platform.

## 3. Deepnote

I love Deepnote. I use it for both data science and machine learning projects. It is an AI-powered cloud Jupyter Notebook filled with features that will make your life easier. You can create projects and invite your team to make real-time changes to both the code and the data. It is a highly collaborative platform that also lets you share applications and reports.

The only drawback of using Deepnote is that you have to pay to access the GPUs.



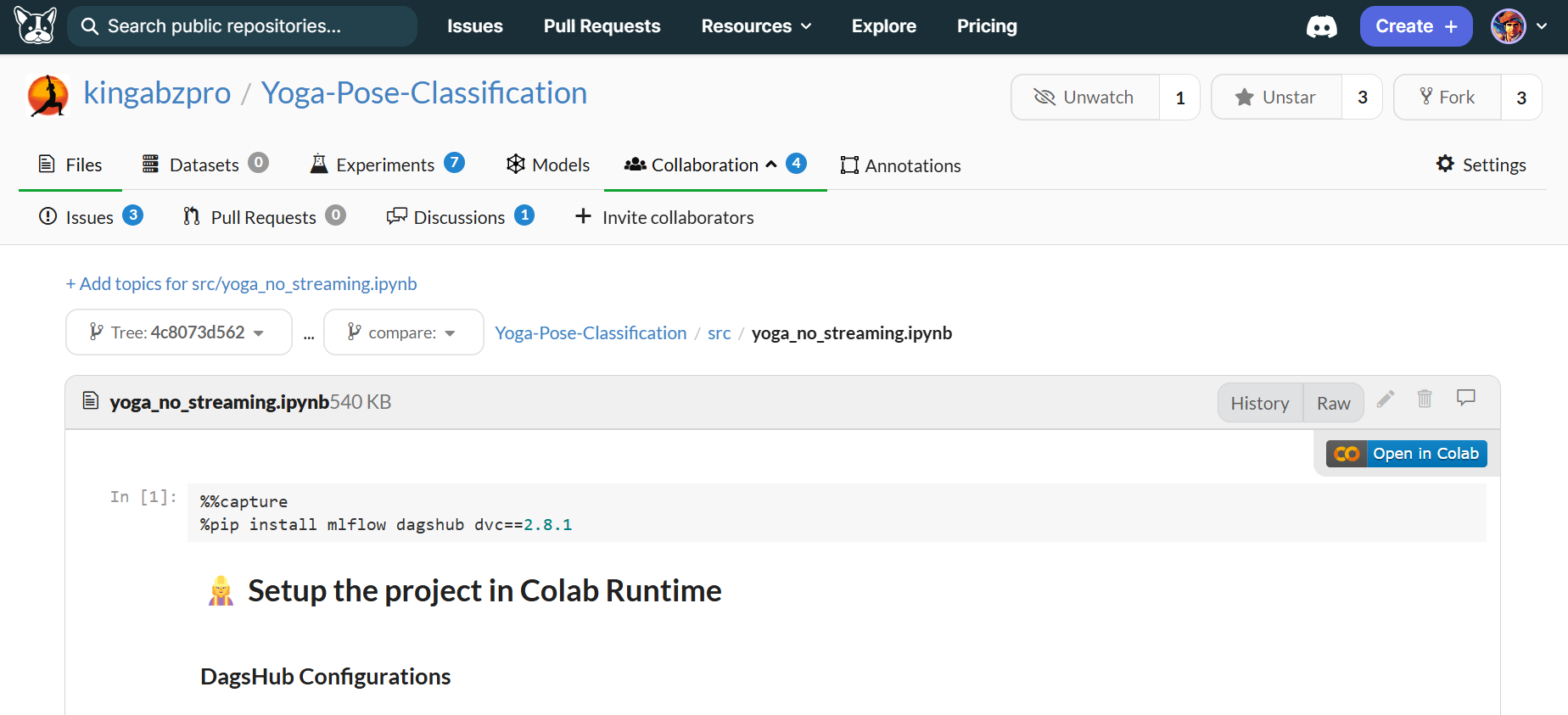
Apart from live collaboration, you can share the project with someone outside the company, comment on specific lines just like in Google Docs, track progress using the history feature, and integrate external data sources.

If you are a beginner, I highly recommend starting with Deepnote to experience its user-friendly yet powerful platform for machine learning project collaboration.

## 4. DagsHub

When I say that DagsHub is a pure machine learning collaboration platform, I mean it. Unlike other platforms that cater to a broader range of developers and data professionals, DagsHub is specifically designed to meet the unique needs of machine learning engineers.

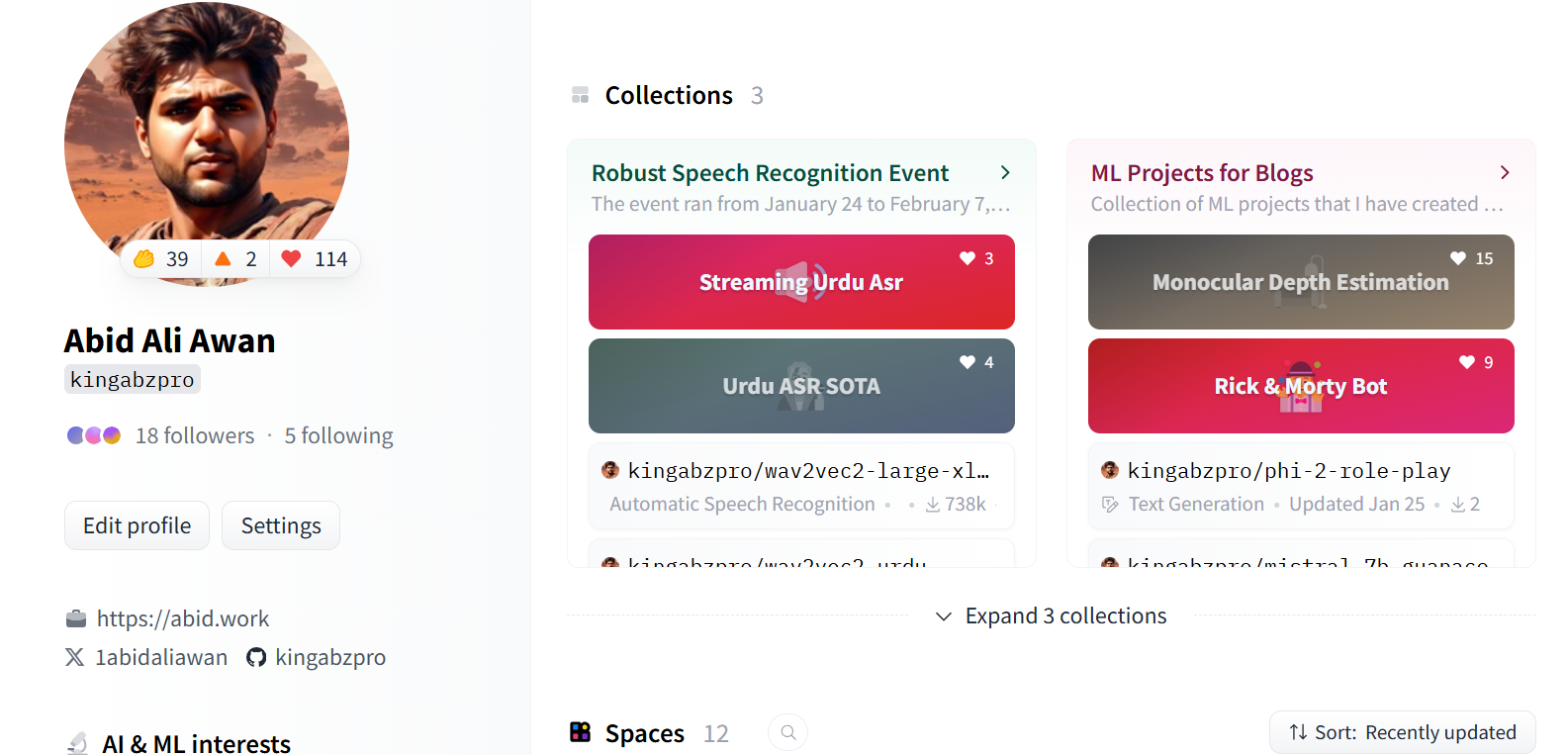
DagsHub provides a platform for sharing code, data, models, and metadata, simplifying collaboration on machine learning projects and boosting productivity. Moreover, it comes with free MLFlow integration, allowing you to track your experiments, save your models, and even serve them with ease.



From data annotation to model serving, DagsHub has got you covered. It includes inline comments, GitHub-style discussion features, webhooks, and external integrations with AWS, GCP, and Azure storage solutions.

## 5. Hugging Face

Hugging Face is a machine learning platform that every beginner and professional should consider. It primarily allows you to share models and datasets, but it also comes with additional features like serving models, deploying machine learning applications, writing posts, commenting on projects, and more. You can use either Git or the Hugging Face API to access and share files.



Slowly, Hugging Face is becoming an all-in-one platform for machine learning engineers. It enables users to share their thoughts like on social media, share models and datasets, serve models, and even offers enterprise-level solutions such as providing GPU support for deployed models.

So, why is it placed at the bottom? Despite its simplicity in collaboration, apart from using Git or the website, your team cannot collaborate extensively on the project, and there are limitations on many aspects.

## Conclusion

Whether you are a beginner just starting out or a seasoned professional, choosing the right platform can significantly enhance your workflow and productivity.

We have explored five exceptional platforms—Kaggle, GitHub, Deepnote, DagsHub, and Hugging Face—each offering unique features tailored to the needs of machine learning engineers. From real-time collaboration and integrated development environments to version control and model serving, these platforms provide a wide range of tools to support your projects.