School of Computing Science
University of Glasgow

MLOps Infrastructure Terraform Scripts

These Terraform scripts automate setting up cloud infrastructure on GCP for an end-to-end machine learning pipeline using AutoML Vision.

Contents

The following resources are provisioned:
- New GCP Project
- Enablement of required APIs like AutoML, Vertex AI
- Service account for AutoML access
- Cloud Storage bucket for images
- Vertex AI dataset for AutoML

The `main.tf` file defines and configures these resources.
`variables.tf` contains input variables like region, account names, API list etc. to customize the deployment.

Usage

1. Update `variables.tf` as needed for your environment.
2. Run `terraform init` to initialize the directory.
3. Run `terraform plan` to preview the changes.
4. Run `terraform apply` to provision the infrastructure.
5. Follow the instruction manual for the rest of the steps to provision MLOps pipeline
6. Run `terraform destroy` when finished to tear down the resources.

Authentication

The scripts use programmatic authentication via the Terraform Google provider. Ensure your local system is authenticated by running `gcloud auth application-default login`.

Customization

- The AutoML model type can be changed by updating the Vertex AI dataset metadata schema URI.
- Additional APIs can be enabled by adding to the `required_apis` list.
- Resource names and properties can be adapted by modifying the variables.

Documentation

For detailed comments explaining each resource, refer to `main.tf`.
The Terraform Google Provider documentation can provide more information on configuring
GCP resources.