Session 2:
Productionising NLP Model - I

SESSION 2: PRODUCTIONISING NLP MODEL - I

- Gain an understanding of the use-case and its business context
- Recap the solution notebook
- Gain an understanding of S3 and CodeCommit
- Clone and understand the cloned repositories
- Make changes to the notebook

Use case

BHARATFIN

- Credit card / Prepaid card
- Bank account services
- Mortgages/loans

BUSINESS CONTEXT

- The company has scaled 10x over 2 years
- Customers raise tickets for services
- They need to be resolved quickly
- Automation of tagging the tickets is essential for handling scale
- Manual errors may occur while assigning tickets

BUSINESS GOAL

KPI	Current Value	Expected Value
% of tickets resolved on time	60%	80%

AUTOMATIC TICKET CLASSIFICATION

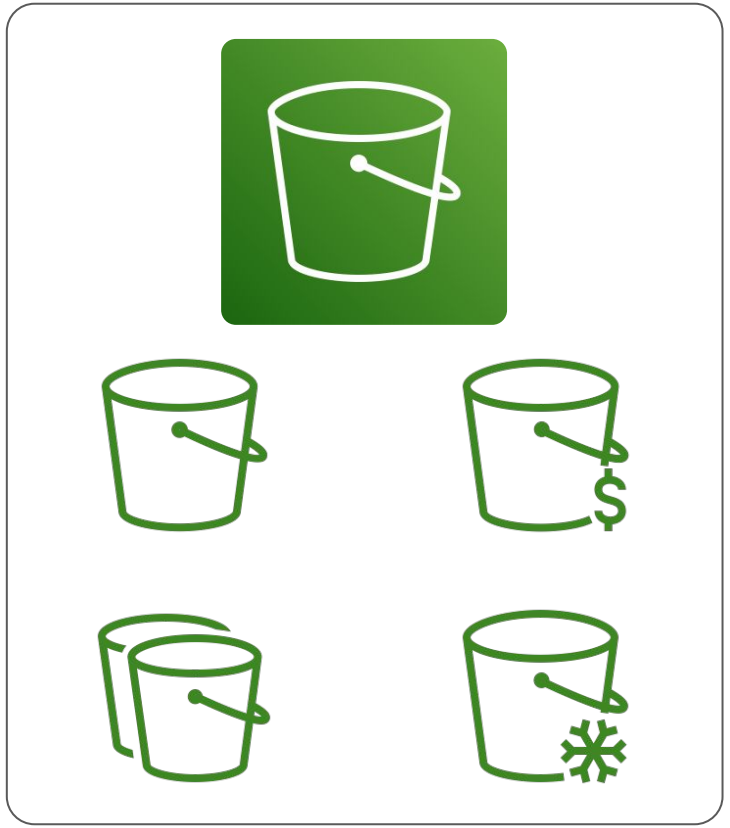
- The customer support team has approached the data science team for automatic ticket classification
- These customer complaints are unstructured text data.
- We will be using an NLP model to classify tickets.

WHY MLOPS-BASED SOLUTION?

- Reduce lag in model development and deployment
- Experiment tracking
- Monitoring tickets will help in reducing incorrect classifications
- Continuous training is required to avoid model decay over time
- **Real-time classification:** The company receives large-volume tickets throughout the day. Hence, it needs to create an API for running the model whenever a ticket arrives.

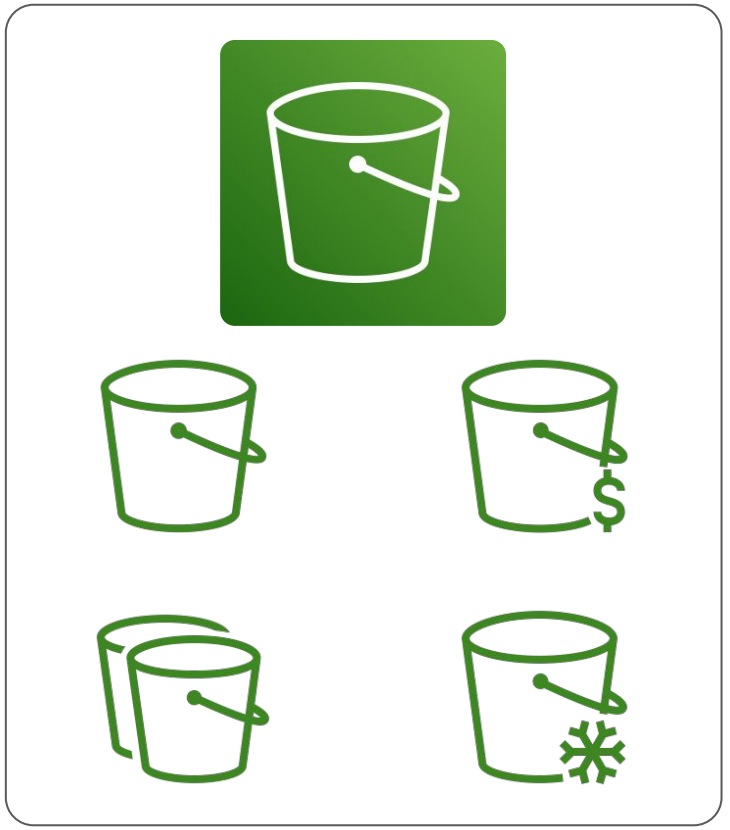
SIMPLE STORAGE SERVICE (S3)

- **Object storage service** allows to store and retrieve any amount of data at any time from anywhere
- Uses
  - Back up and store data
  - Archive data
  - Host static websites
  - Acts as intermediate storage for other AWS Services
- Advantages
  - Designed for **99.999999999%** (total: eleven 9s) data durability
  - **99.99%** available: Possible owing to data replication in three availability zones
  - Highly Scalable



KEY CONCEPTS

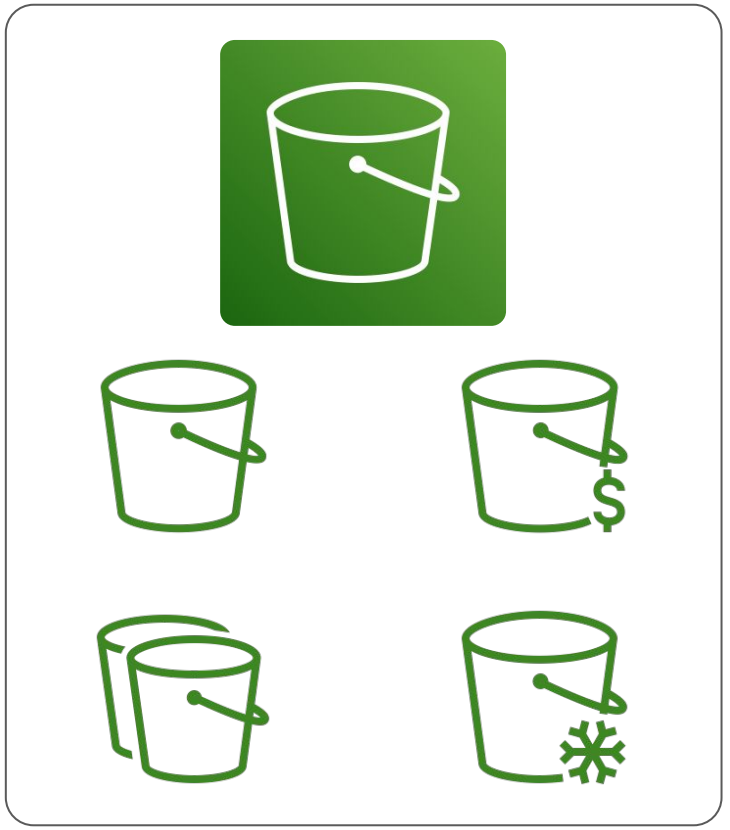
- Buckets
  - Contain objects
  - Unique name globally
- Keys: Unique object identifier
- Objects
  - Object Data
  - Meta Data
- Versioning



WHY S3

S3 will be used to store the following:

- Data sets
- Tar files for model
- Code Artifacts
- Preprocessing and evaluating scripts



CODE COMMIT

- Provides a version control service
- Stores and manages assets in the cloud
- Hosts private Git Repositories
- Eliminates the need for you to manage your source control system
- Supports the standard functionality of Git

HOW CODECOMMIT WORKS?

- Provides a console
  - Easily creates repositories
  - Lists existing repositories and branches
- Finds information about a repository and clones it
- Creates a local repository, makes changes and pushes it
- Uses a command line or GUI-based editor

SESSION 2: SUMMARY

- Gained an understanding of the use case and its business context
- Recap of the solution notebook
- Gain an understanding of S3 and CodeCommit
- Clone and understand the cloned repositories
- Make changes to the notebook