Building Applications with Vector Databases

Use Cases

Semantic Search

Retrieval Augmented Generation (RAG)
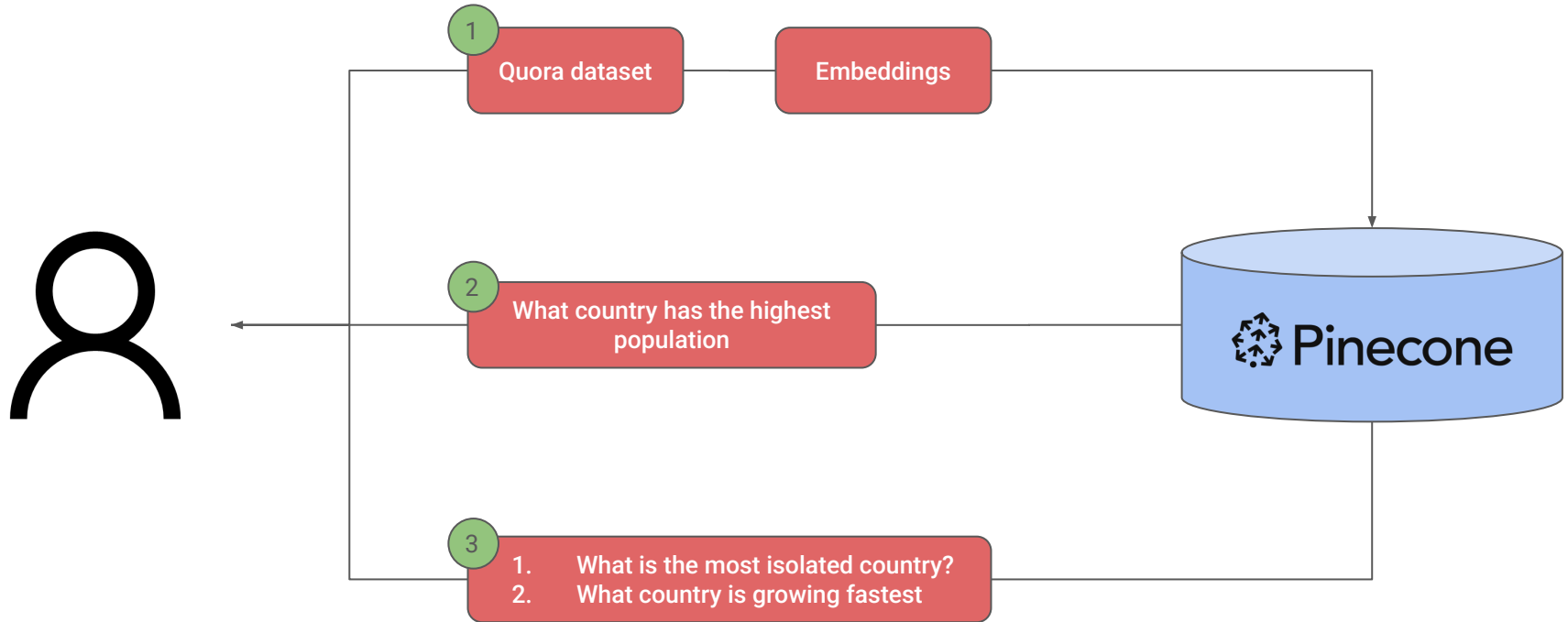
Recommender Systems

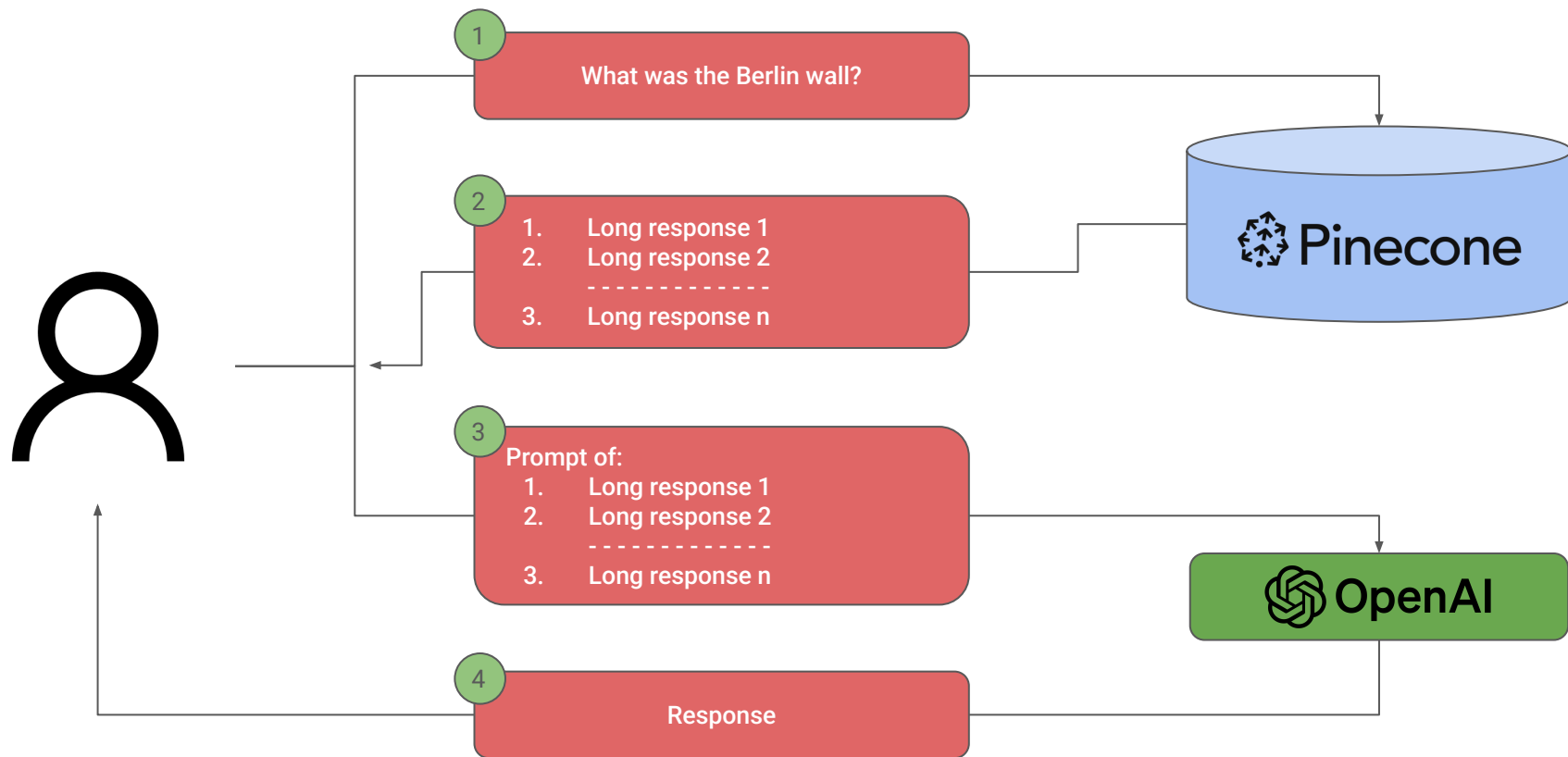Hybrid Search

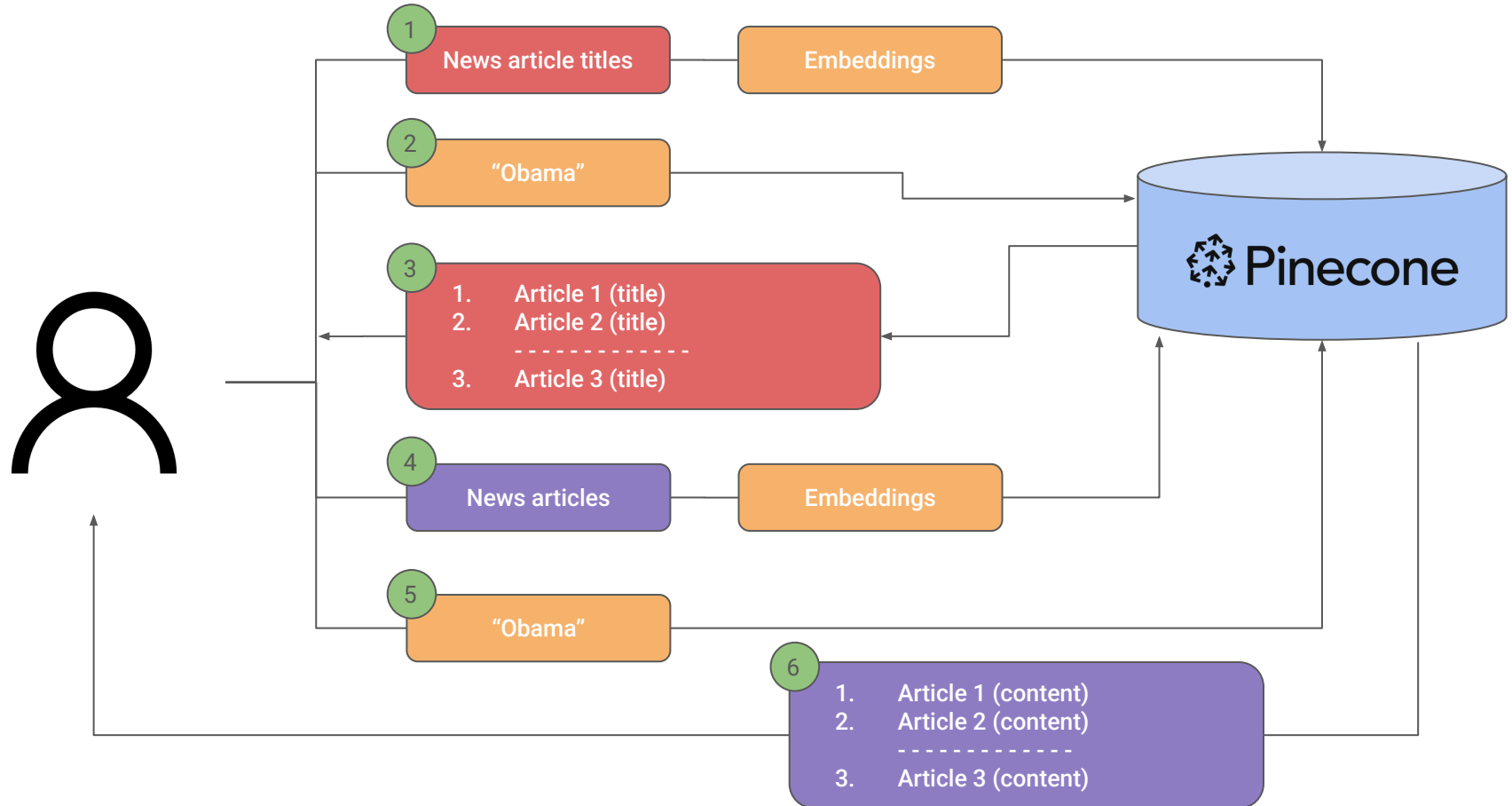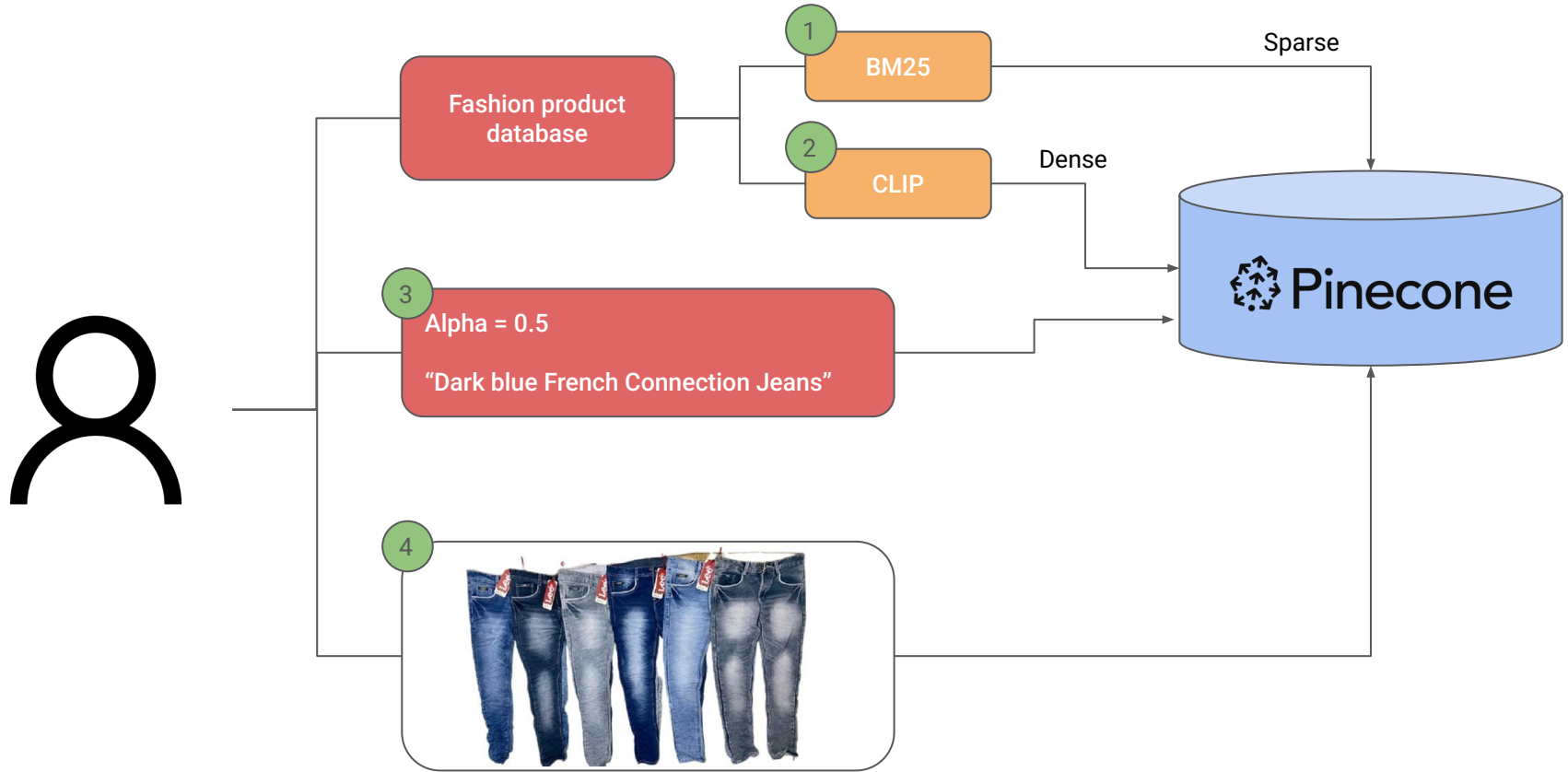Facial Similarity Search

Anomaly Detection

Semantic Search

Retrieval Augmented Generation

Recommendation System

Hybrid Search

Sparse Vector using BM25

BM25 is a ranking function that is used to retrieve text by estimating the relative importance of terms in the text of the search query.
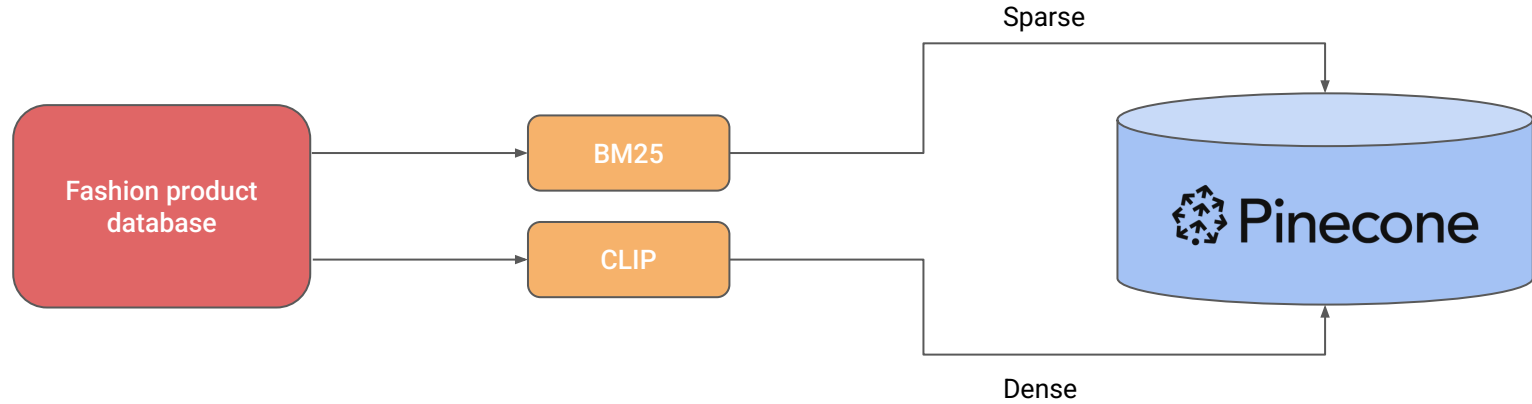
It is calculated based on the number of documents in the data corpus and the word frequency across all relevant documents

CLIP

CLIP (Contrastive Language-Image Pre-Training) is a
neural network trained on a variety of (images, text)
pairs.

It maps an image to text/caption to describe the image

Sparse-Dense Vector

Generate sparse and dense vectors for the full dataset

Upsert these vectors with its metadata to Pinecone

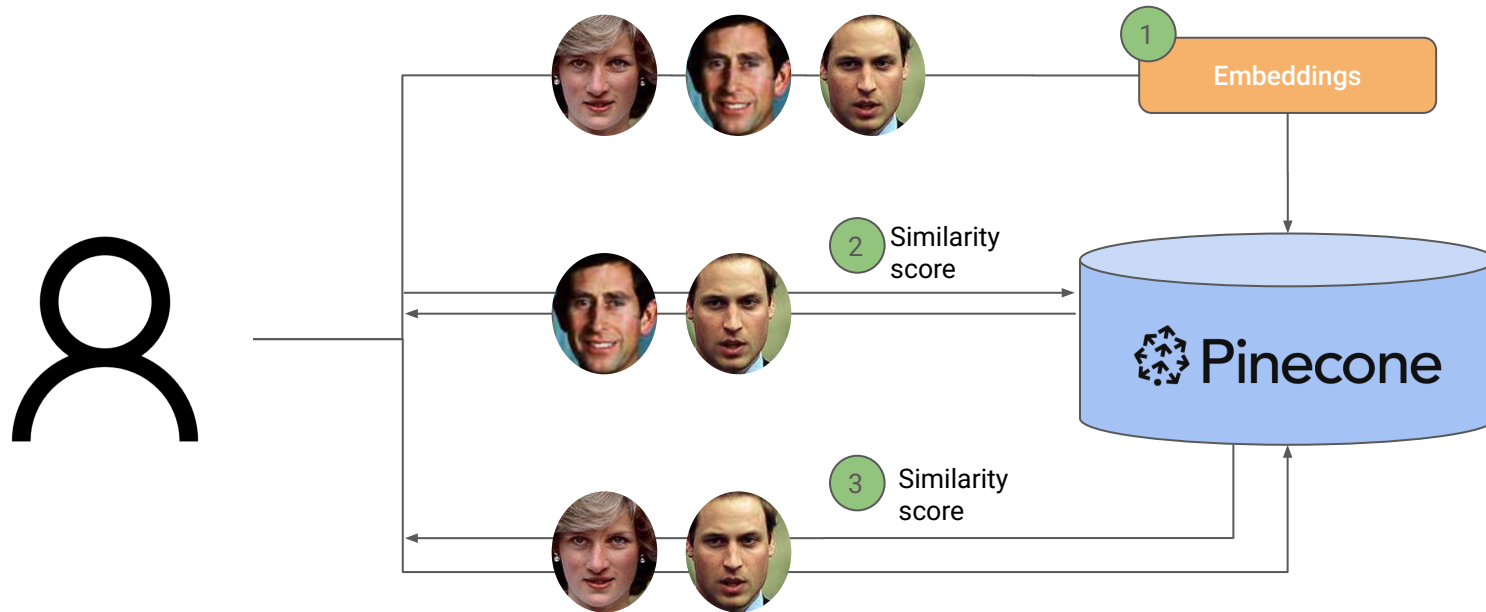# Facial Similarity Search

Mother

Father

Child

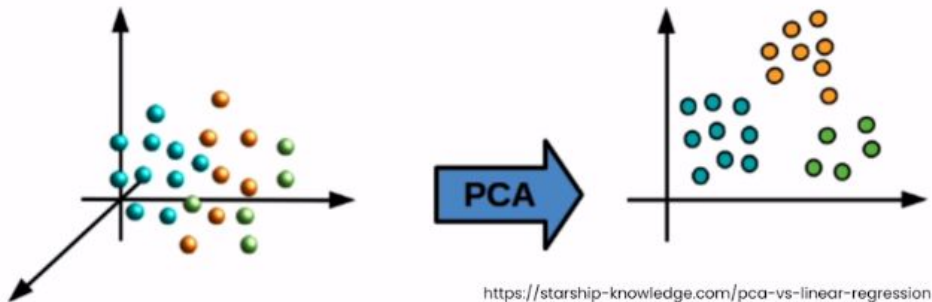# Facial Similarity Search

Visualize a data of Images

Two steps

1. **PCA:** Preliminary step to reduce the dimension of our data of embeddings
2. **T-SNE plot:** to plot the data after PCA

Principal Component Analysis

PCA, or principal components analysis, reduces the dimensionality of any data. Such as, the data is embedding of many images

PCA can help us in visualizing and plotting our data by generating a new set of variables that can represent all the embeddings in our data

https://starship-knowledge.com/pca-vs-linear-regression

t-SNE Plot

t-SNE, or t-Distributed Stochastic Neighbour Embedding, is a tool to visualize a high dimensional data like data of multiple image embeddings.

It works by grouping similar data points together based on their similarities in the higher dimensional space.

This technique is highly effective for understanding complex datasets. It allows for identifying patterns and clusters and uncovering connections between data points that may not be immediately obvious in higher dimensions.
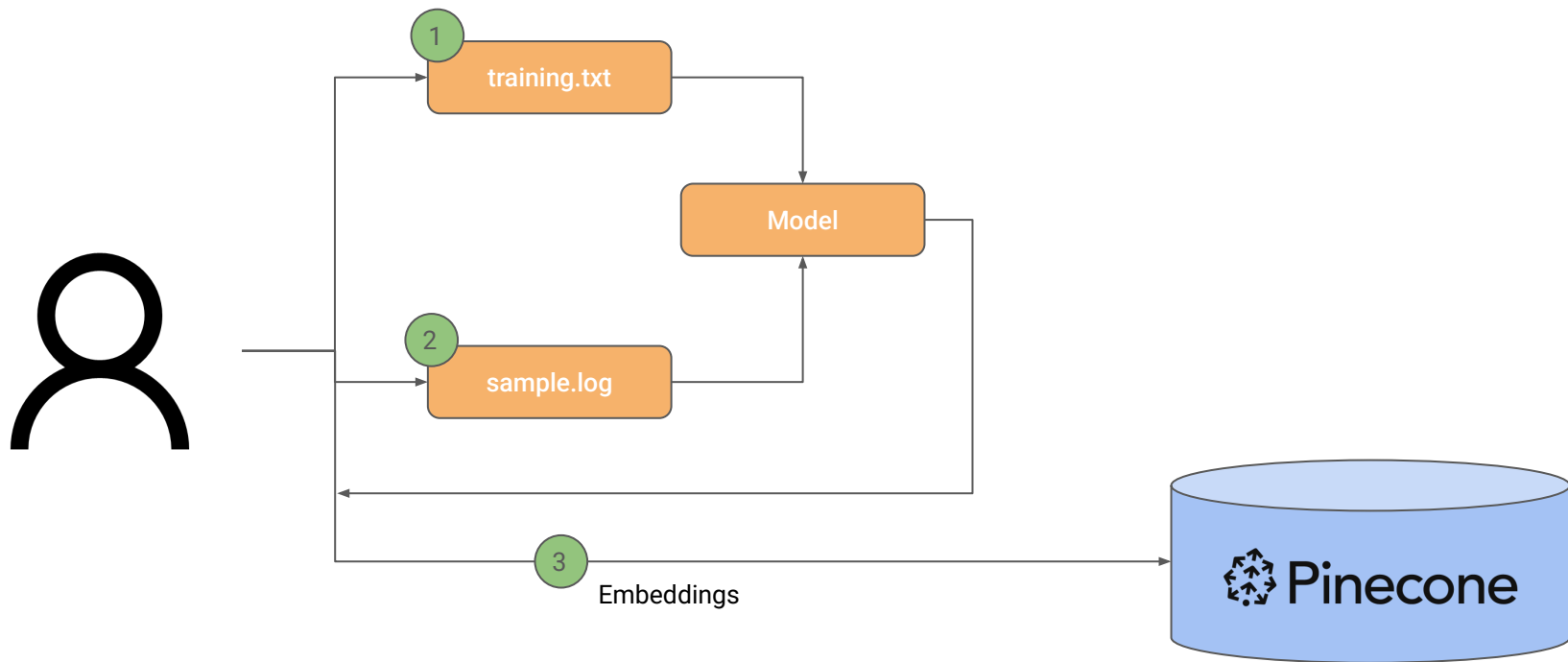
Visualize a data of Images

Two steps

1. **PCA:** Preliminary step to reduce the dimension of our data of embeddings
2. **T-SNE plot:** to plot the data after PCA

A key parameter of t-SNE's is "Perplexity", which represents the number of neighbors each data point has in the higher dimensional space

Anomaly Detection

THANK YOU