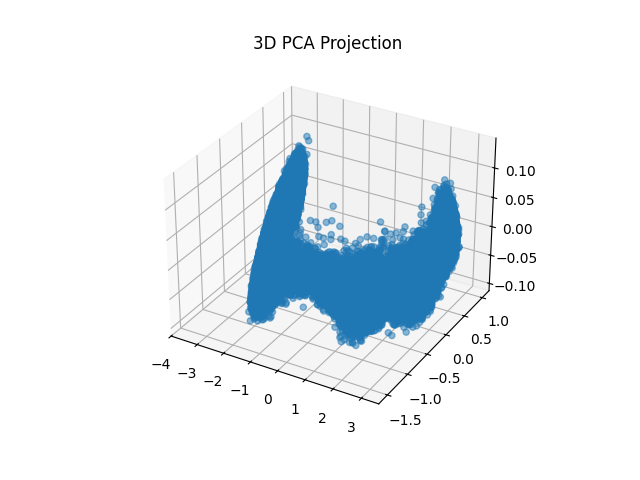
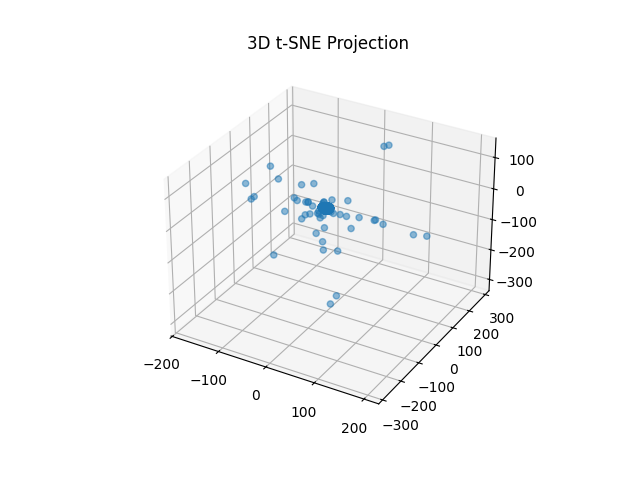
**Analysis of Model Performance and Embedding Visualization**

****

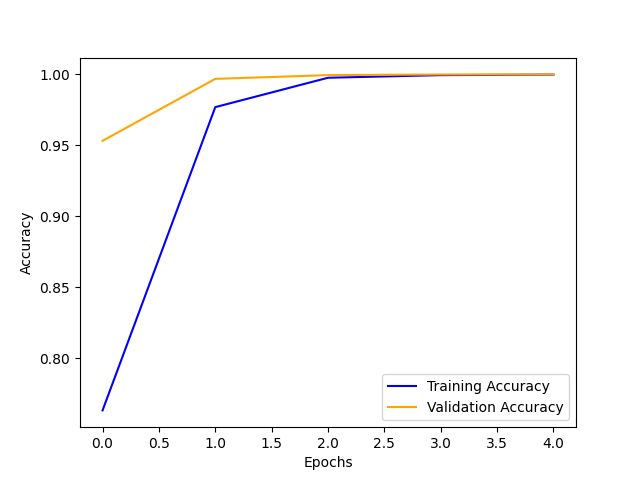
**1. 3D PCA Projection**

* **Description**: This image shows a 3D Principal Component Analysis (PCA) projection of the model's embeddings.
* **Purpose**: PCA is commonly used to reduce the dimensionality of data, making it easier to visualize complex, high-dimensional datasets. Here, PCA captures the main variance in the embeddings in three dimensions, providing insight into clustering and separation in the data.
* **Observations**:
  + There appears to be a concentration of points around specific regions, with some clear spread along certain axes.
  + This suggests that the embeddings may have some distinct clusters, indicating that the model is learning separable features or patterns in the data.
* **Interpretation**: The clustering observed might suggest that different groups of similar data points are positioned closer in the embedding space. This is useful if the embeddings will later be used for tasks like clustering, classification, or retrieval.

****

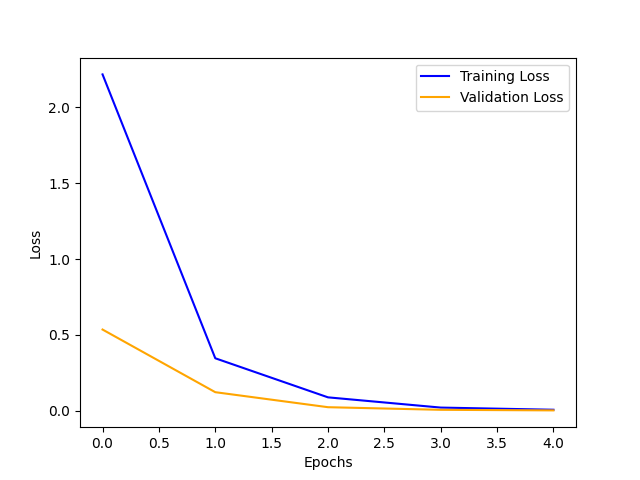
**2. 3D t-SNE Projection**

* **Description**: This 3D t-SNE (t-Distributed Stochastic Neighbor Embedding) projection visualizes the same embeddings with a different dimensionality reduction technique.
* **Purpose**: t-SNE is well-suited for visualizing high-dimensional data in a low-dimensional space (often 2D or 3D) by preserving the local structure. It often highlights clusters more clearly than PCA.
* **Observations**:
  + The points are more dispersed than in the PCA projection, which could mean that t-SNE has managed to reveal finer-grained local clusters within the data.
  + There is a distinct central cluster with some points scattered around, suggesting stronger local relationships among certain data points.
* **Interpretation**: The clustering here indicates that the embeddings may capture nuanced relationships in the data, which is valuable for downstream tasks like classification or question-answering. t-SNE’s visualization reinforces that the model has likely learned meaningful relationships.



**3. Training and Validation Accuracy Over Epochs**

* **Description**: This line plot shows the training and validation accuracy across five epochs.
* **Observations**:
  + Both training and validation accuracy increase rapidly in the initial epochs, with training accuracy reaching close to 100% by the end.
  + Validation accuracy also remains high, indicating minimal overfitting.
* **Interpretation**: The high accuracy achieved for both training and validation sets suggests that the model is fitting well without significant overfitting, which is a positive indicator of model generalization. The stability in the accuracy curves toward the end implies the model has converged.



**4. Training and Validation Loss Over Epochs**

* **Description**: This plot tracks the training and validation loss across the same five epochs.
* **Observations**:
  + Both training and validation loss decrease sharply in the first few epochs, with training loss approaching zero.
  + Validation loss stabilizes at a low value, which aligns with the high validation accuracy observed.
* **Interpretation**: The rapid reduction in loss suggests efficient learning during the early stages of training. The low and stable validation loss reinforces the idea that the model has not overfit and is learning relevant features in the data.