The **goal** of Franceschi et al. method is to extract meaningful low-dimensional representations from **multivariate time series (MTS)** using **unsupervised learning**. These representations can later be used for various downstream tasks such as clustering, forecasting, and anomaly detection.
The core idea is to use **neural networks** to encode high-dimensional MTS into low-dimensional latent representations.

- **Encoder:**
  - **Convolutional layers** are applied to the input MTS to capture **local temporal patterns** across multiple variables.
  - An **LSTM** network processes the extracted temporal features from the convolutional layers to capture **long-term dependencies** between variables.
- **Decoder:**
  - The decoder attempts to **reconstruct** the original time series from the encoded representations using a **sequence-to-sequence (Seq2Seq)** prediction model.
  - This reconstruction forces the encoder to learn representations that summarize the important features of the input data.

Once the model generates **latent representations** (low-dimensional embeddings) for the multivariate time series (MTS) data, these representations are passed to the **K-Means algorithm** for clustering. The idea is that these embeddings summarize the relevant temporal patterns and dependencies in the original time series, making them suitable for clustering. Since the original time series can be very high-dimensional, the latent embeddings serve as a **compressed and meaningful summary** of the data. These embeddings are now suitable inputs for clustering algorithms like **K-Means**.

Applying K-Means:

- The **K-Means algorithm** is applied on the set of latent representations.
- K-Means assigns each time series (via its latent representation) to one of **K clusters** by minimizing the distance between points and the cluster centroids in the embedding space.

Clustering Output:

- The output is a set of **clusters** where each cluster contains time series that exhibit similar temporal patterns.
- The quality of clustering is directly related to the **quality of the learned embeddings**—better embeddings lead to more meaningful clusters.

However when we tested this method it did not perform well for the pattern recognition. The figure below visualizes the centroids for the 4 discovered clusters and highlights the nearest cluster for each part of the original time series. As the figure demonstrates, portions of the time series which are similar to each other are assigned to different clusters. Interpreting the time series in terms of the clusters is challenging because of these inherent differences and the inability to map the cluster centroid directly onto portions of the original data.