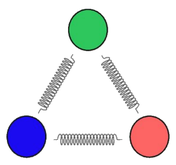
CONTRASTIVE FORECASTING

Latent-Space Time Series Prediction Using Contrastive Divergence

Contrastive Forecasting is an unsupervised time series prediction method that operates in latent space by aligning predicted future states with actual ones while contrasting against negative examples.

INTRODUCTION

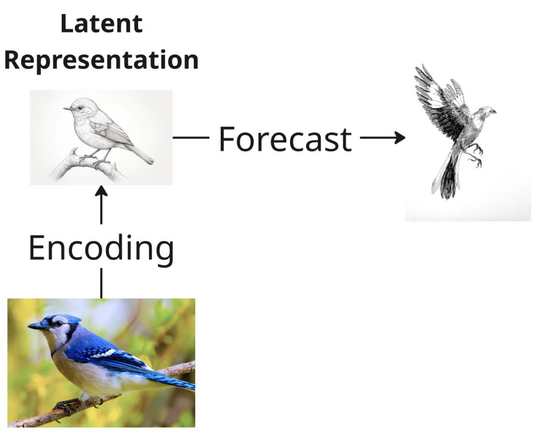
CONTRASTIVE DIVERGENCE



Contrastive divergence is a principle that “pulls” positive (similar) samples closer in the representation space and “pushes” negative (dissimilar) samples farther apart.

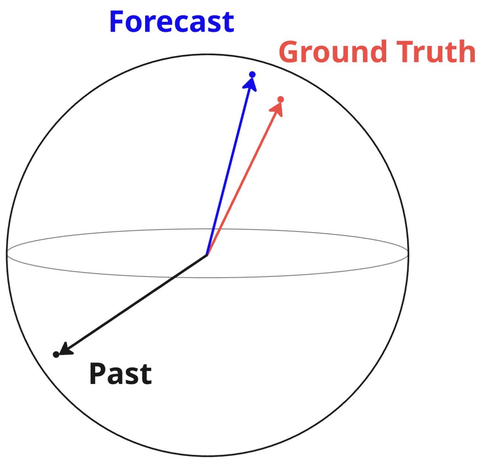
$$\mathcal{L} = -\sum_{i=1}^{N} \log \left(\frac{\exp(\text{sim}(x_i, x_i^+)/\tau)}{\sum_{x_j^- \in \mathcal{N}_i} \exp(\text{sim}(x_i, x_j^-)/\tau)} \right)$$

Latent representation



Forecast in the latent space

Forecasted latent representation of a temporal slice and ground truth encoding of the same temporal slice are positives. Encoding of a past temporal slice is a negative.



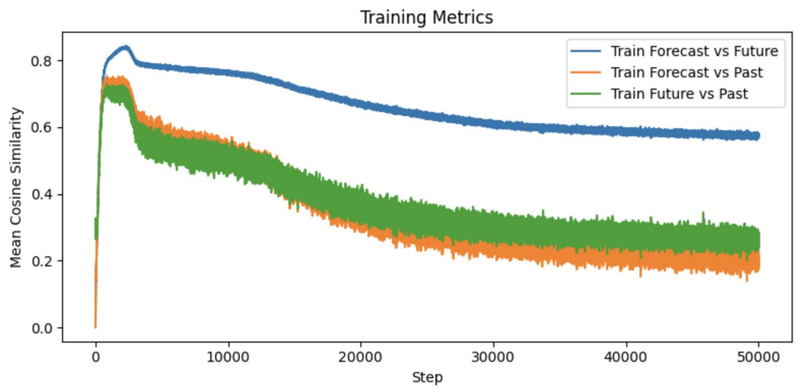
RESEARCH

LOSS FUNCTION

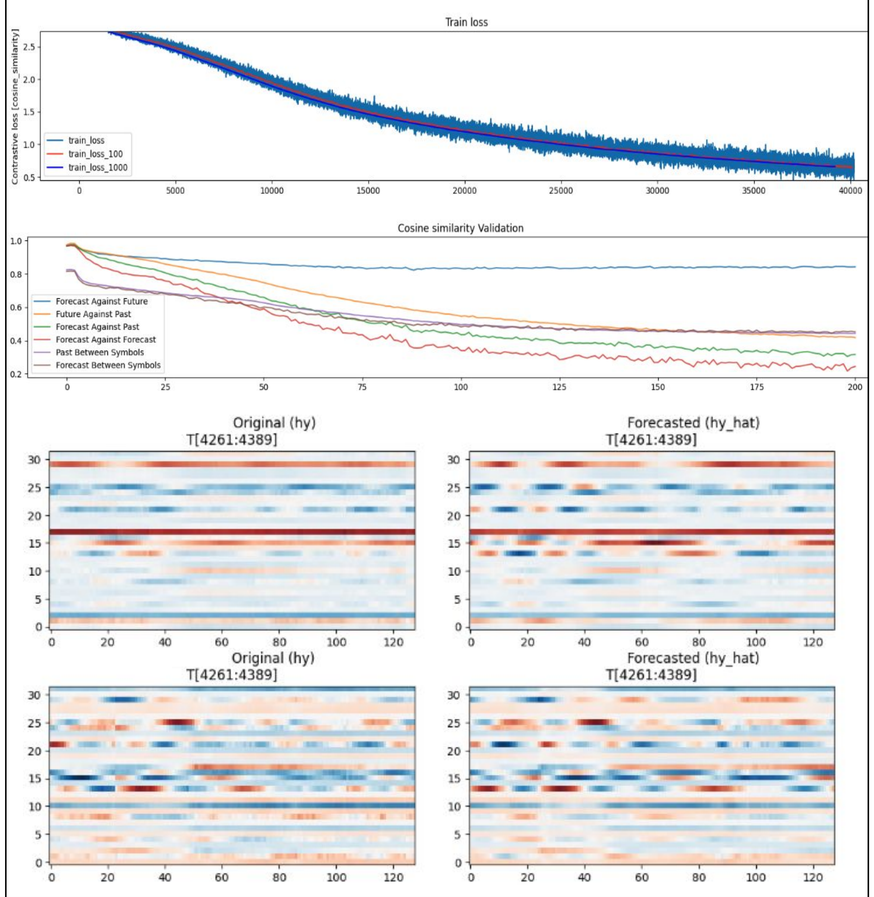
$h_{t+1}^i \leftrightarrow \hat{h}_t^i$	positif future and forecast
$h_t^i \nleftrightarrow h_{t+1}^j$	negatif past and future
$h_t^i \nleftrightarrow \hat{h}_t^j$	negatif past and forecast
$h_t^i \underset{i \neq j}{\nleftrightarrow} h_t^j$	negatif between past channels
$\hat{h}_t^i \underset{i \neq j}{\nleftrightarrow} \hat{h}_t^j$	negatif between forecast channels
$\hat{h}_t^i \nleftrightarrow \hat{h}_{t+1}^i$	negatif consecutive forecasts *
$h_{t+1}^i \underset{i \neq j}{\nleftrightarrow} \hat{h}_t^j$	negatif future and forecast between channels *
$\hat{h}_t^i \nleftrightarrow \hat{h}_{t+1}^j$	negatif consecutive forecast between channels *

$$\mathcal{L} = -\frac{1}{TB} \sum_{T,B} \log \left(\frac{\sum_i \exp\left(\frac{\text{sim}(x_i, x_i^+)}{\tau}\right)}{\sum_y \sum_{i,j} \exp\left(\frac{\text{sim}(x_i, y_j^-)}{\tau}\right)} \right)$$

Synthetic Data



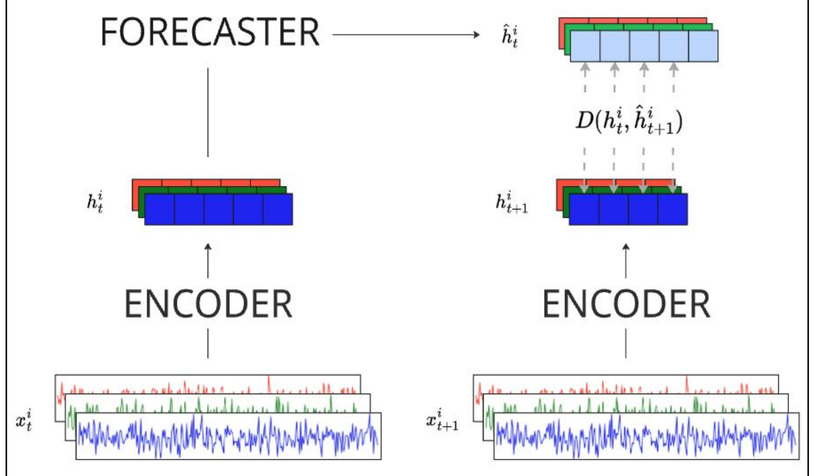
Real Data



IMPLEMENTATION

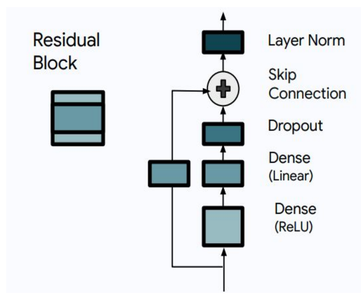
CONTRASTIVE FORECAST

We slice multivariate time series into patches. We send each patch to a latent encoding. We trains a model to predict the next patch encoding by using a contrastive loss that pulls the forecasted encoding closer to the actual next patch while pushing it away from the original patch and other channels.



ARCHITECTURE

The forecaster is a RWKV or Transformer network. The encoder is a ResidualBlock as in TimeFM.



Das, A., Kong, W., Leach, A., Mathur, S. K., Sen, R., & Yu, R. (2023). Long-term Forecasting with TiDE: Time-series Dense Encoder. Transactions on Machine Learning Research.

CONTACT INFO

jeremy.cochoy@gmail.com
@jeremycochoy [LinkedIn, Telegram, GitHub]

We are recruiting a Master’s/PhD student to join this project.