

CRUNCH Seminars at Brown, Division of Applied Mathematics

Friday - January 22, 2021

Adversarial Sparse Transformer for Time Series Forecasting

Jeremy Chen

The paper addresses the accumulation error problem when conducting multi-step forecasting (i.e., this is essentially when we append the model's own output to the real values and use it to forecast subsequent time steps). The paper also addresses creating more diverse forecasts with multiple ranges of values. To address these problems the author purposes using GANs. It is one of the first articles that I have seen that describes using GANs for forecasting. The GANs are used as a method of regularizing multi-step time series predictions. They work in conjunction with a sparse attention mechanism that utilizes an ent-max activation function instead of Softmax.

This allows the network to better learn long range dependencies in time steps and, in particular, which steps aren't important at all.