Machine Learning for Time Series Data (Seminar)

Preliminary Meeting (IN2107, IN4874).

Dr. Tim Januschowski (Zalando)
Dr. Jan Gasthaus (AWS AI Labs)

Summer Term 2022

Team

- Dr. Tim Januschowski (Zalando)
- Dr. Jan Gasthaus (AWS AI Labs)

This is a seminar for **Master** students!
Main prerequisite: Machine Learning (IN2064)

Website

`https://www.in.tum.de/daml/lehre/sommersemester-2022/seminar-time-series/`

Over to you: introductions

- Who are you?
- Why are you participating/what do you want to take away?

Goals

1. Learn about and explore **state-of-the-art** research in ML for time series
2. **Analyze and criticize** recent publications or **dive deep** into a method and explore extensions/improvements
3. Improve your **scientific writing**
4. Participate in a **review process** akin to international conferences
5. Improve your **presentation skills**

Tasks

1. Read **seed research papers** (provided by us)
2. **Choose** either
   2.1 **Snowball research:** identify and read additional papers related to the seed papers (via references to/from the paper, relevant keywords)
   2.2 **Deep dive:** Experiment with the code released with the paper; extend/improve the method/code, run experiments, and analyze the results
3. Summarize your findings, criticism, and research ideas in a **short paper** (4 pages, double column)
4. Write **reviews** of other students work
5. **Present** your work in 25-minute talks + 10 minutes discussion

Grade will be based on **all** parts: Paper, reviews, presentation, and participation in discussion

Snowball Research: Paper Outline

1. Introduce the topic
2. Summarize the research presented in the paper
3. Relate the paper to other work
   - What is different / novel in this paper? Why is it important?
   - Are there other papers that build on this one? How and why?
4. Discuss & critique the paper, suggest improvements
   - Is the solution proposed in the paper sensible? What would be an alternative?
   - Are all claims reasonably supported by arguments and/or experimental evidence? Which experiments are missing? Are baseline methods or commonly-used data sets missing in the experiments?
   - Is there relevant prior work that the authors are not citing?

Tasks: Deep Dive Deliverables

- Pre-requisite: read and understand the paper and decide with supervisor on the direction of extension
- Deliverable (i): code & experiments.
- Deliverable (ii): paper
  - short summary of the area
  - overview of the method
  - presentation & discussion of own extension
  - presentation of results
- Deliverable (iii): talk
- Deliverable (iv): active participation in Seminar

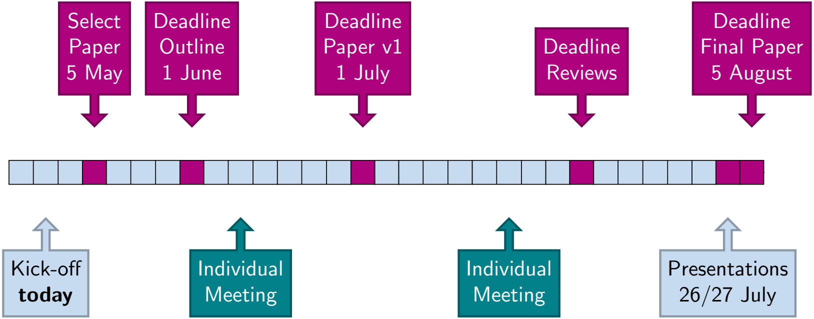
Note: it is possible for one student to choose a paper for a snowball research and another student to choose the same paper for a deep dive. Apart from that, we won't have team work formally given guidance from university based on the difficulty in grading individual contributions in team work.

Review Process

- We'll mimick the paper review of ML conferences

  https://nips.cc/Conferences/2020/PaperInformation/ReviewerGuidelines
- You will review 3-4 papers each if you choose snowball research (for deep dive, no reviews are necessary)
- Goal is to **evaluate** the paper (strengths and weaknesses) and to provide feedback the helps the author to improve the paper
  1. Summary and contributions: Briefly summarize the paper and its contributions
  2. Strengths: Describe the strengths of the work. Typical criteria include: soundness of the claims (theoretical grounding, empirical evaluation), significance and novelty of the contribution, and relevance to the ML community.
  3. Weaknesses: Explain the limitations of this work along the same axes as above.
  4. Correctness: Are the claims and method correct?
  5. Clarity: Is the paper well written?
  6. Additional feedback, comments, suggestions for improvement and questions for the authors

Schedule


Select Paper 5 May
Deadline Outline 1 June
Deadline Paper v1 1 July
Deadline Reviews
Deadline Final Paper 5 August
Kick-off **today**
Individual Meeting
Individual Meeting
Presentations 26/27 July

Further Details

- Individual meetings via biweekly virtual office hours: Wednesday's 9:30 am to 11:30. 30 mins slots. Sign up here: `https://docs.google.com/spreadsheets/d/123MSfXj4Q_rPtCcgNsaIz7tQt5xxFVAexEjPPP5BrTU/`
- Dial-in details to follow.
- Office hours are full? Contact us, we'll try to find a solution.
- You want to choose your own seed paper? Come to office hour on 4 May and we'll discuss & decide. Send the paper to us before hand by email.

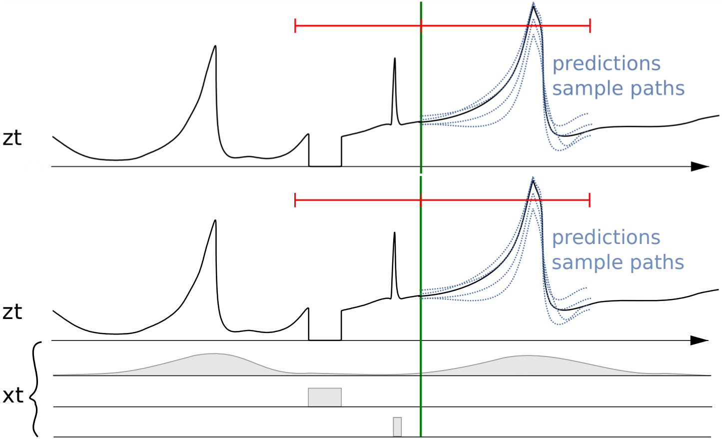
Seminar: 26 & 27 July

- 1.5 days, exact times, schedule and location tbd
- current preference: in person
- first day: full day of presentations (25 mins + 10 mins questions)
- evening first day: dinner (we'll invite you)
- second day: half day of presentations
- active participation means: ask questions in the discussion.

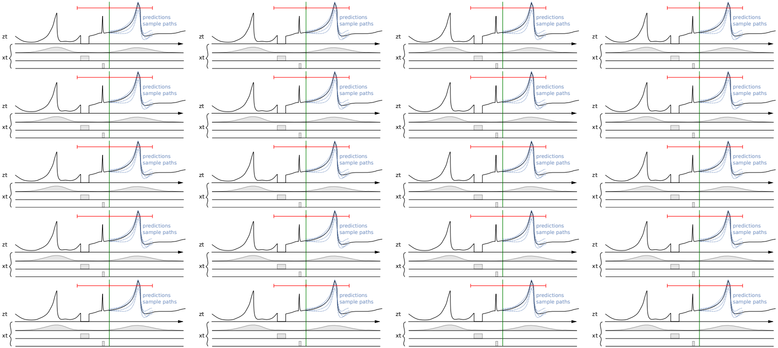
Forecasting Problem Setup

Predict the future behavior of a time series given its past

$$z_1, z_2, \ldots, z_T \Longrightarrow P(z_{T+1}, z_{T+2}, \ldots z_{T+\tau})$$



Typical Setup: panel of *related* time series, exogenous variables.

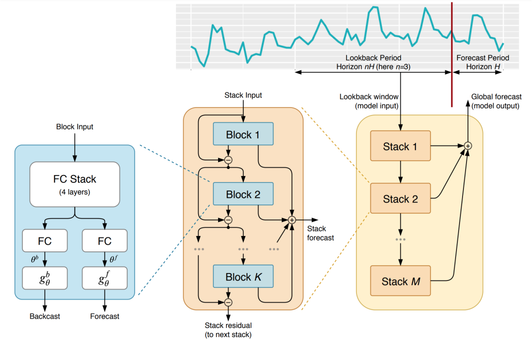


We have a panel of time series $Z = \{\mathbf{z_i}\}_{i \in I}$ and we want to forecast each time series in the panel. Typically $|I|$ is large.

Topic: Point or Probabilistic Forecasting

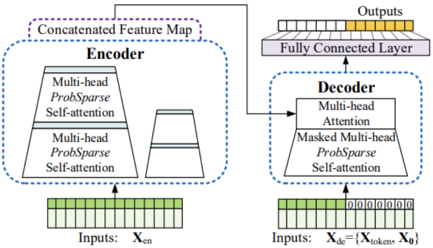
- what are good state of the art models for forecasting?
- how to obtain probabilistic forecasts?
- questions for deep dive: how to go from point to probabilistic forecast?

Seed Paper: N-BEATS



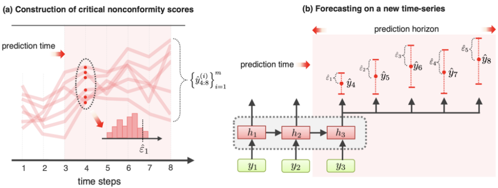
- Generic architecture pure ML model that achieves SOTA on M4
- Only Point Forecasts
- exogenous variables not inclued
- `https://arxiv.org/abs/1905.10437`

Seed Paper: Informer



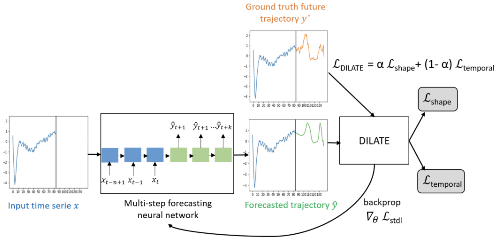
- Transformer Architecture with a couple of Time Series Twists
- Only Point Forecasts
- exogenous variables not inclued
- `https://ojs.aaai.org/index.php/AAAI/article/view/17325`

Seed Paper: Conformal Time Series Forecasting



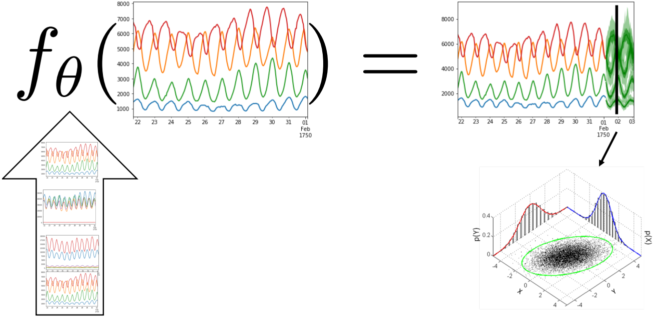
- conformal prediction: obtaining probabilistic predictions with very limited assumptions
- use errors on calibration set to obtain prediction intervals
- `https://proceedings.neurips.cc/paper/2021/file/312f1ba2a72318edaaa995a67835fad5-Paper.pdf`

Seed Paper: Dilate



- introduces a new loss for forecasting
- uses a shape and a temporal term
- point forecasts only
- `https://proceedings.neurips.cc/paper/2019/file/466accbac9a66b805ba50e42ad715740-Paper.pdf`

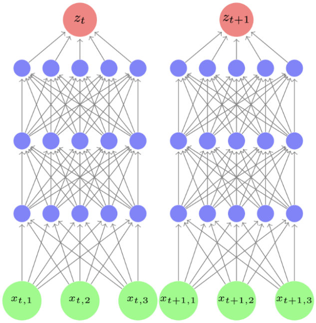
Topic: Multivariate Forecasting Model

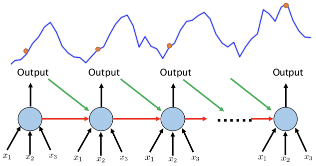


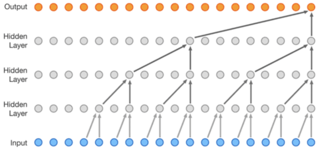
Seed Papers

- https://arxiv.org/pdf/2002.06103.pdf
  - combination of transformer & normalizing flows
  - capture dependency structure as part of the flows
- https://papers.nips.cc/paper/2020/file/cdf6581cb7aca4b7e19ef136c6e601a5
  Paper.pdf
  - use GNNs to capture dependency structure in the spectral domain
  - uses Graph and Discrete Fourier Transforms
  - only point forecasts

Architectures for Sequence Data







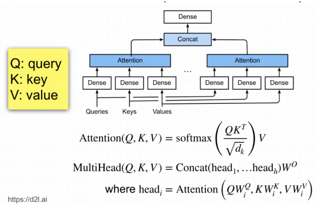


Figure credit: ICML 2019 tutorial on Attention in Deep Learning by Alex Smola and Aston Zhang

Architectures for Sequence Data

RNN DeepAR: Probabilistic forecasting with autoregressive recurrent networks
David Salinas, Valentin Flunkert Jan Gasthaus, Tim Januschowski
International Journal of Forecasting, Volume 37, Issue 3, July-September 2021, Pages 1302-1303

CNN An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling
Shaojie Bai, J. Zico Kolter, Vladlen Koltun
arXiv Preprint, arXiv:1803.01271

Att Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting
Bryan Lim, Sercan O. Arik, Nicolas Loeff, Tomas Pfister
International Journal of Forecasting, Volume 37, Issue 4, October-December 2021, Pages 1748-1764

Hierarchical Time Series

Multivariate time series with a hierarchical structure

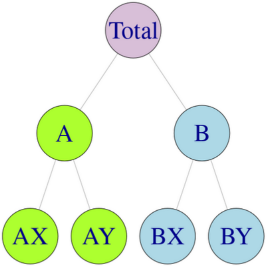
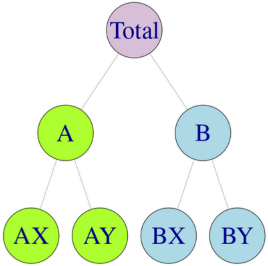


Figure: An example hierarchy (Hyndman et al., 2018).

- Daily sales of all laptops
- Daily sales of a particular brand of laptops
- Daily sales of a given brand in a particular location

Hierarchical Time Series

Hierarchical time series satisfies linear aggregation constraints

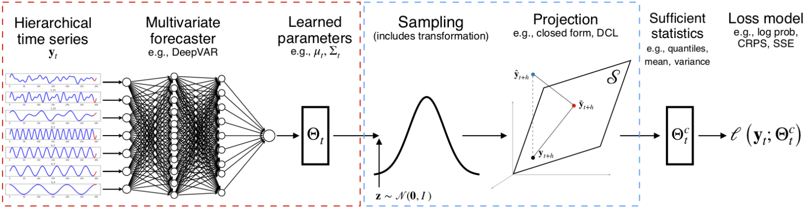


$$\mathbf{y}_t = S\mathbf{b}_t$$

$$\begin{bmatrix} y_t \\ y_{A,t} \\ y_{B,t} \\ y_{AX,t} \\ y_{AY,t} \\ y_{BX,t} \\ y_{BY,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_{AX,t} \\ y_{AY,t} \\ y_{BX,t} \\ y_{BY,t} \end{bmatrix}$$

Seed Paper: End-to-End Learning for Hierarchical Forecasting

Idea: A single *trainable* model



Ensure differentiability

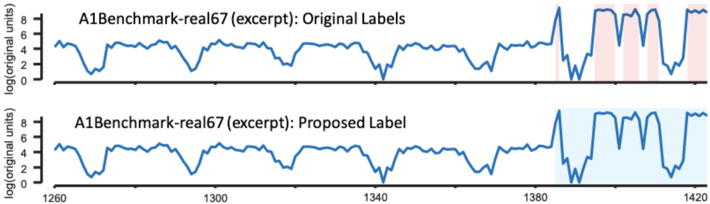
- Sampling: reparameterization trick
- Projection: multiplication with a pre-determined matrix.

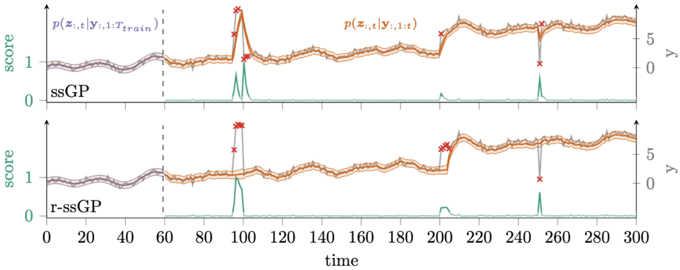
http://proceedings.mlr.press/v139/rangapuram21a.html

Seed Paper: SHARQ

- Idea: use regularization techniques along the hierarchy
- Use quantile regression for probabilistic forecasts
- `http://proceedings.mlr.press/v130/han21a/han21a.pdf`

Time Series Anomaly Detection

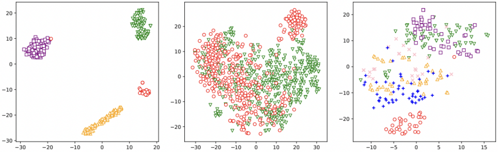




Time Series Anomaly Detection

1. Timeseries Anomaly Detection using Temporal Hierarchical One-Class Network
   Lifeng Shen, Zhuocong Li, James Kwok
   NeurIPS 2020
2. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery
   Thomas Schlegl, Philipp Seeböck, Sebastian M. Waldstein, Ursula Schmidt-Erfurth, Georg Langs
   International conference on information processing in medical imaging. 2017
3. Robust random cut forest based anomaly detection on streams
   Sudipto Guha, Nina Mishra, Gourav Roy, Okke Schrijvers
   ICML 2016

Time Series Representation Learning



(a) DiatomSizeReduction. (b) FordB. (c) OSULeaf.

- Map time series into high-dimensional embedding space independent of particular ML task
- enable multiple tasks like classification, forecasting
- questions that papers answer:
  - what architectures?
  - what losses?
  - what tasks do we want to specialize in?

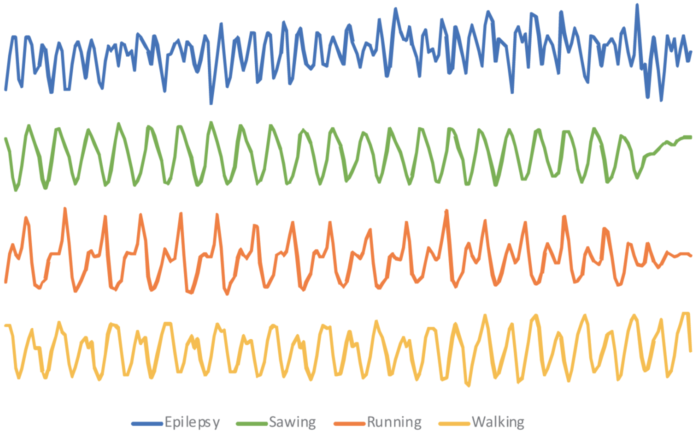
Seed Papers

- Transformers and auto-regressive denoising
  https://arxiv.org/pdf/2010.02803.pdf
- CNNs and contrastive losses
  https://arxiv.org/abs/1901.10738
- Temporal Neighbourhood encoding
  https://openreview.net/forum?id=8qDwejCuCN
- use seasonal/trend decomposition (classic idea) for representation learning https://openreview.net/forum?id=PilZY3omXV2

Sample deep dive questions

- try a different loss on given architecture (or vice versa)
- attempt a new task with the representation

Time Series Classification



Time Series Classification

1. MINIROCKET: A Very Fast (Almost) Deterministic Transform for Time Series Classification
   Angus Dempster, Daniel F. Schmidt, Geoffrey I. Webb
   KDD 2021
2. InceptionTime: Finding AlexNet for time series classification
   Hassan I. Fawaz et al.
   Data Mining and Knowledge Discovery, Volume 34, Issue 6, Nov 2020, pp. 1936-1962
3. Time Series Classification with HIVE-COTE: The Hierarchical Vote Collective of Transformation-Based Ensembles
   Jason Lines, Sarah Taylor, Anthony Bagnall
   ACM Transactions on Knowledge Discovery from Data, Volume 12, Issue 5, October 2018

Additional Topics

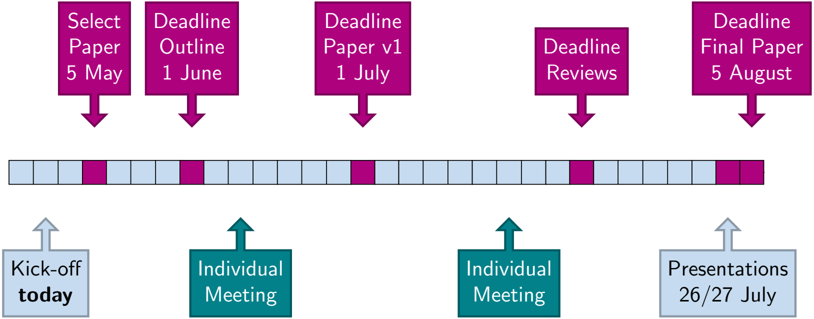
1. **Explainability for time series models**
   Explaining Time Series Predictions with Dynamic Masks
   Jonathan Crabbé, Mihaela Van Der Schaar
   ICML 2021
2. **Causality**
   Causal Forecasting: Generalization Bounds for Autoregressive Models
   Leena C. Vankadara, Philipp M. Faller, Lenon Minorics, Debarghya Ghoshdastidar, Dominik Janzing
   arXiv Preprint, arXiv:2111.09831

Paper Selection

- sign-up here: `https://docs.google.com/spreadsheets/d/123MSfXj4Q_rPtCcgNsaIz7tQt5xxFVAexEjPPP5BrTU/`
- only sign up for one paper. If you're favorite paper is already taken either choose a different one or put yourself on the waiting list. We'll find a solution.
- deep dive & snowball research by two different folks on the same paper is possible
- till 5 May
- bring-your-own paper: possible, but email us first and come to office hour.

Questions & Answers

Schedule


Select Paper 5 May
Deadline Outline 1 June
Deadline Paper v1 1 July
Deadline Reviews
Deadline Final Paper 5 August
Kick-off **today**
Individual Meeting
Individual Meeting
Presentations 26/27 July