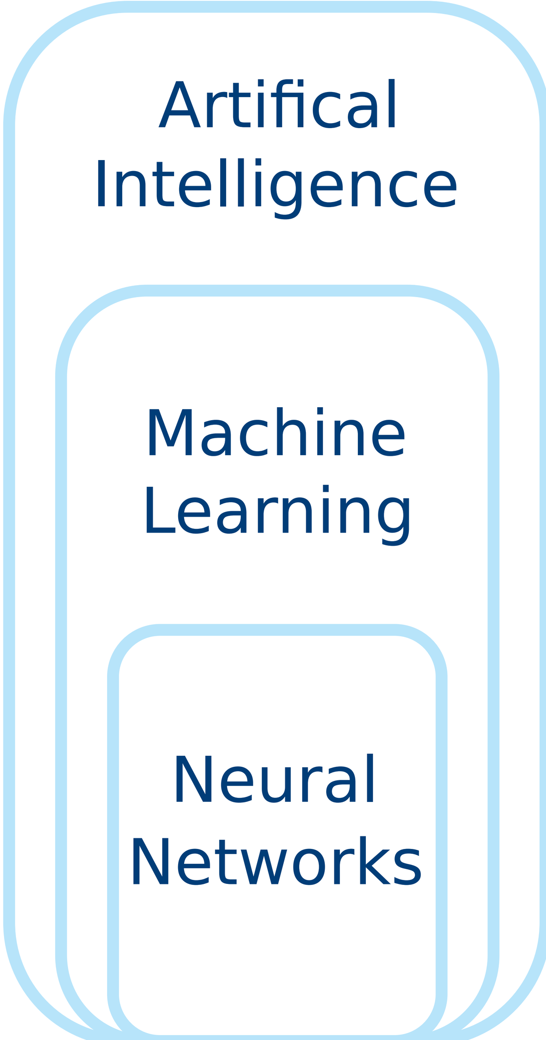
Explainable Machine Learning for Climate Prediction

Jonathan Ortved Melcher, Andrea Vang, André Düsterhus, Hjalte Danielsen Sørup & Jens Hesselbjerg Christensen

What is Machine Learning?

Machine learning has become ubiquitous in climate science, however, it is still a new field with many scientists being in doubt of the definitions and terminology. To clarify; machine learning (ML) is a subset of artificial intelligence (AI), where AI can cover all decisions made by machines, ML requires a statistical basis used to analyse data, and the ability to generalise. Many climate scientists have already employed ML methods such as EOF. In this thesis, I will use a subset of ML called neural networks[1]. These methods mimic the brain's structure and have shown significant predictive power, especially in non-linear cases.

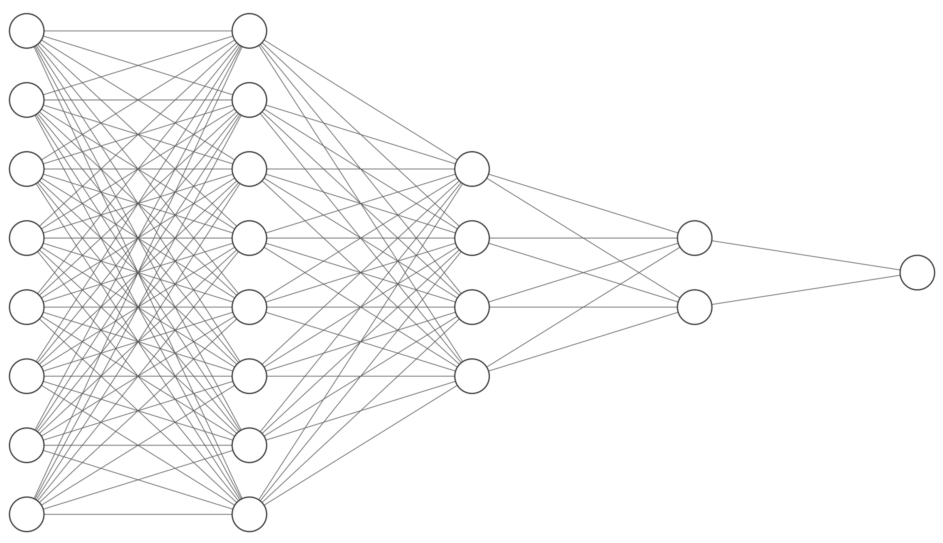


Peering into the Black-Box

Explainable machine learning is a technique developed to understand what inputs have the most effect on the predicted outcome allowing us to understand the connections seen by the ML method and gain a deeper understanding of the statistical relations in the data. In this thesis, I will use layerwise relevance propagation (LRP), with relevance being a measure of how important the neural connection was for the outcome. The relevance is backpropagated using purposely chosen rules. Here the activation function for the layer, special layers such as convolutions, max-pooling and rectified units all determine the rule. Cuttoff rules removing relevance if below a threshold are utilized to denoise the output. The main idea in this thesis is to build on[2] Here they used FFNN to predict heatwaves and then LRP to find that a specific pattern of SSTs led to European heatwaves. In this thesis, the new idea is the introduction of CNNs and long short-term memory networks, which encode spatial and temporal information into the data and are therefore better suited for working climate data that are inherently spatiotemporal related.

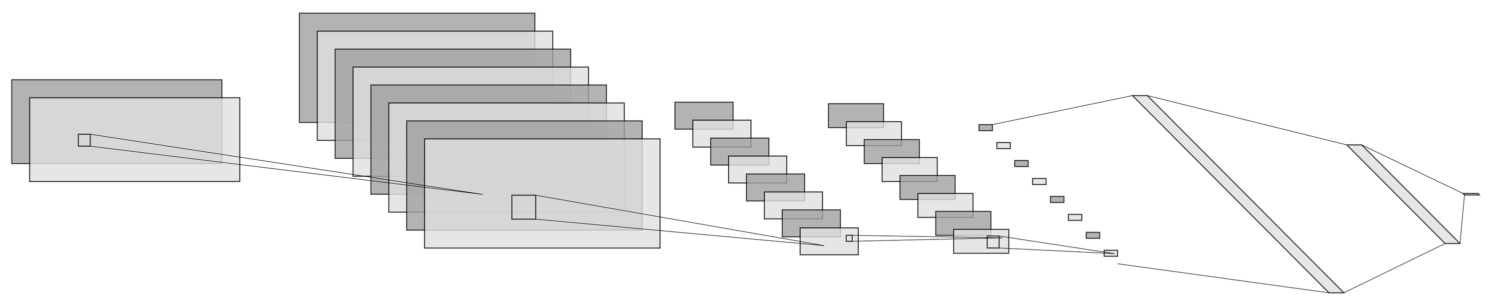
Feed Forward Neural Network (FFNN)

A netowrk with fully connected layers feeding information forward leading to a prediction in the end. This network does not encode spatial or temporal relations



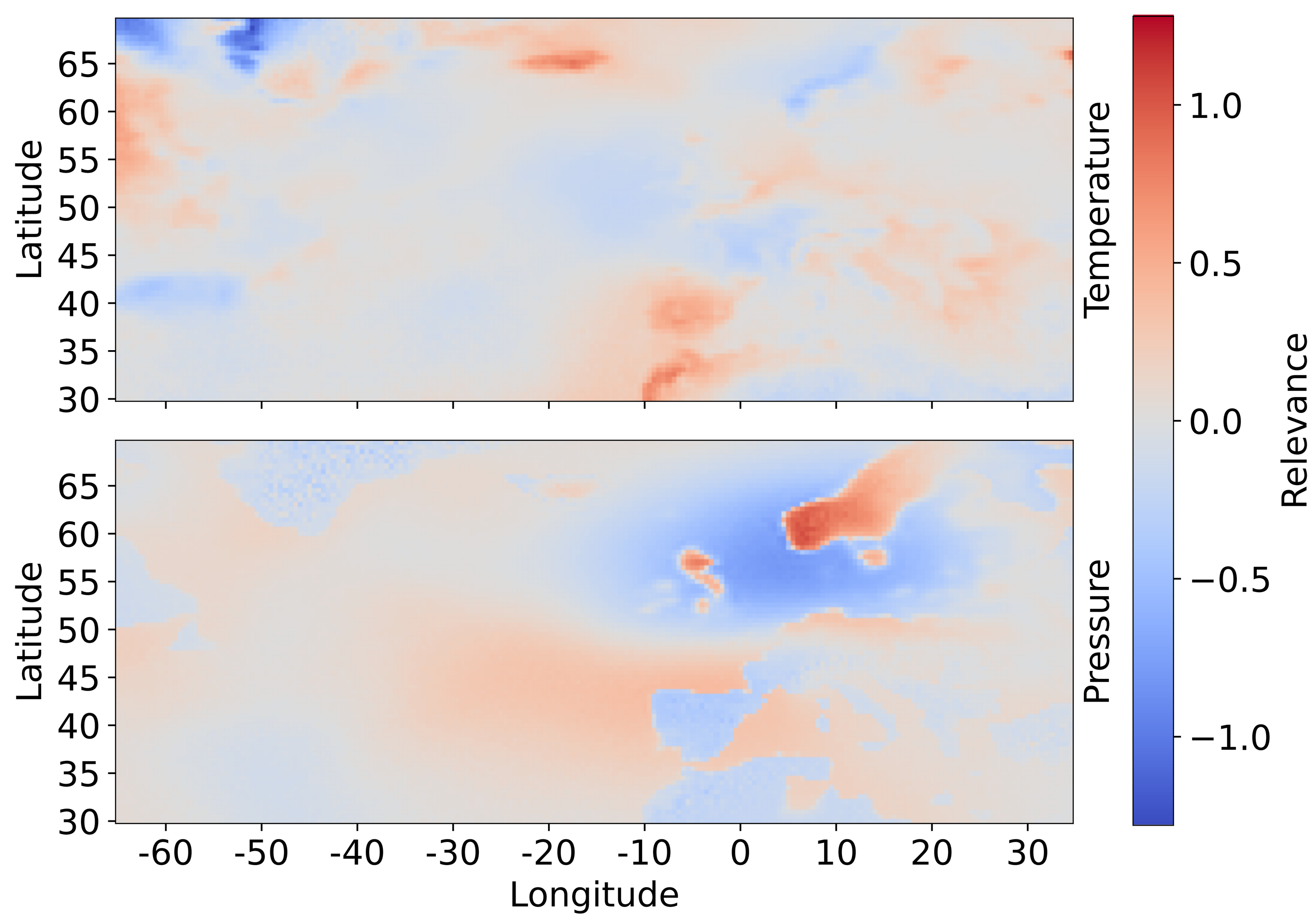
Convolutional Neural Network (CNN)

CNNs utilize convolutions and subsampling to encode the spatial structures seen in data regardsless of their specific position. This was specifcaly developed to recognise objects in picure no mater their position.



Precipitation Predictions

My thesis will try to predict whether a given monthly surface temperature and pressure field will lead to more or less precipitation than the climatological mean. As predictive fields, a large section of the North Atlantic and Europe has been chosen, while for the target a mean over a smaller area in the North Sea baltic region is made. These variables are all normalised on a monthly basis and within each ensemble. The classification will then be inspected using LRP to see what patterns are important for the prediction. This has already been done for ERA53 data as seen to the right. Showing a strong signal with pressure in Bergen, and little signal in temperature. ERA5 does have a disadvantage as the ensembles are not independent and there are only 10 from 1940-2023 giving not enough months to work with thereby leading to overfitting.



CMIP6 and Time-lag

The project will now work on CMIP6 data to avoid overfitting. This contains 19 ensemble members with years from 1850-2014, giving plenty of data points compared to variables. Preliminary results using pointwise correlation between mean precipitation and pressure and temperature, show a similiar pricture to the ERA5 data with a strong signal from pressure in Bergen, and spurious correlations in temperature. From here a classification into 30/40/30 percentiles will be done and LRP utilized to find patterns that lead to each classification. Lastly, an attempt to implement LSTMs to incorporate temporal information will also be made but an increased understanding of LRP in the context of LSTMs is needed[4].

References

[1] Quistgaard, Thea. 2024. "A Basic Introduction to AI and ML."Beobide-Arsuaga,

[2] Goratz, André Düsterhus, Wolfgang A. Müller, Elizabeth A. Barnes, and Johanna Baehr. 2023. "Spring Regional Sea Surface Temperatures as a Precursor of European Summer Heatwaves." Geophysical Research Letters.

[3] Hersbach, Hans, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, et al. 2020. "The ERA5 Global Reanalysis." Quarterly Journal of the Royal Meteorological Society.

[4] Montavon, Grégoire, Sebastian Lapuschkin, Alexander Binder, Wojciech Samek, and Klaus-Robert Müller. 2017. "Explaining Nonlinear Classification Decisions with Deep Taylor Decomposition." Pattern Recognition.