SATNAC 2022

THE FUTURE OF CLOUD AND ITS IMPACT ON INDUSTRIES AND TOTAL HUMAN EXPERIENCE

Multi-Layer Perceptron for Channel State Information Estimation: Design Considerations

Presenter: Andrew Oosthuizen

Co-authors: Prof MH Davel, Prof ASJ Helberg

Hosted by Telkom

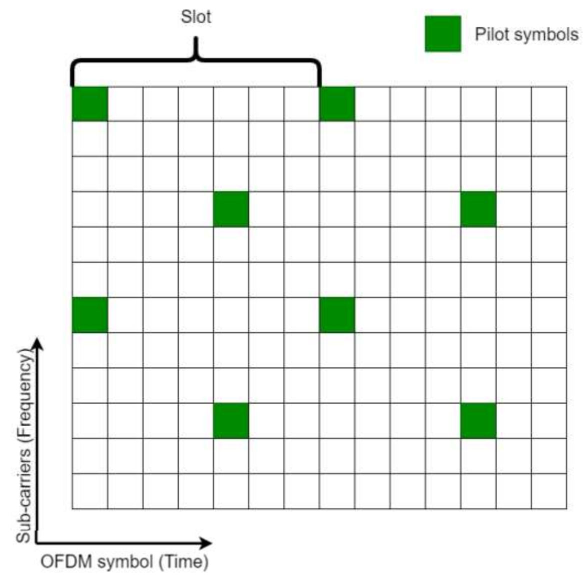
Overview

- Introduction
- Background
- Experimental setup
- Analysis and results
- Conclusion

Channel state information (CSI) estimation

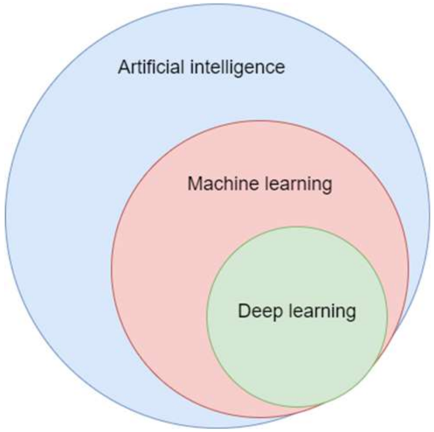
- Channel state information estimation is done to approximate the channel conditions so that received signals can be appropriately equalised.
- Pilot data compares the received data to the expected data.
- This process is complex and computationally exhaustive. Thus, less optimal methods are commonly used for speed.

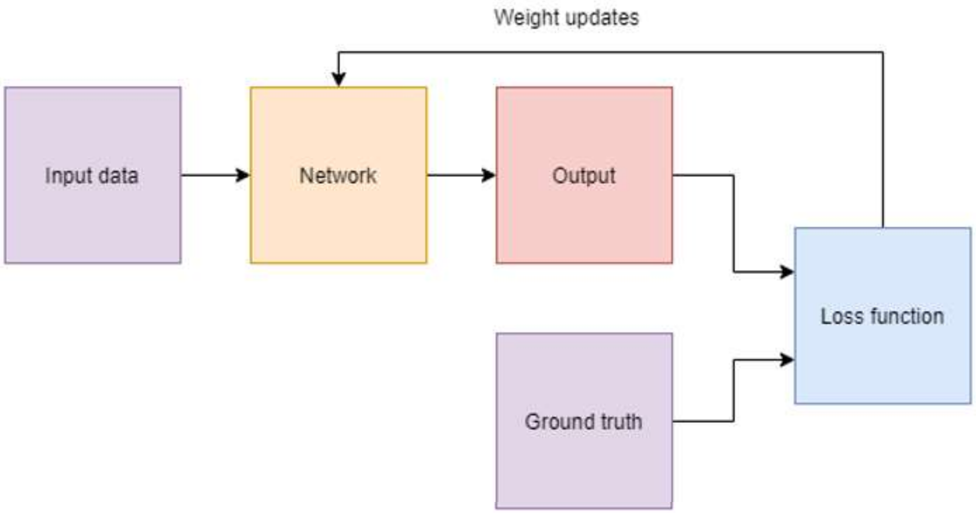
$$H(k) = \begin{bmatrix} h_{1,1}(k) & \cdots & h_{1,m}(k) \\ \vdots & \ddots & \vdots \\ h_{n,1}(k) & \cdots & h_{n,m}(k) \end{bmatrix}$$



Deep learning

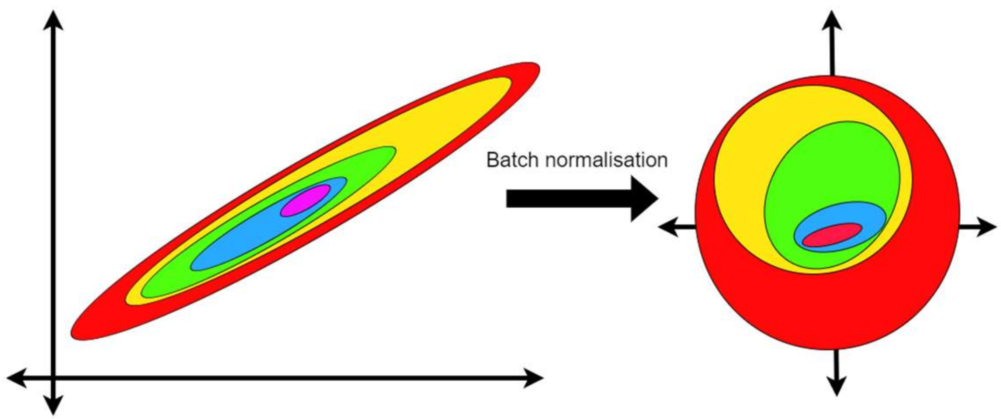
- What is deep learning?
- How are deep learning models trained?
- What is an MLP?
- Why is deep learning attractive in the telecommunications domain?





Batch normalisation

- Batch normalisation is a technique used to normalise activations within the neural network using batches.
- This has the effect of:
  - Speeding up training
  - Acts as reguliser
  - Decreases the importance of initial weights

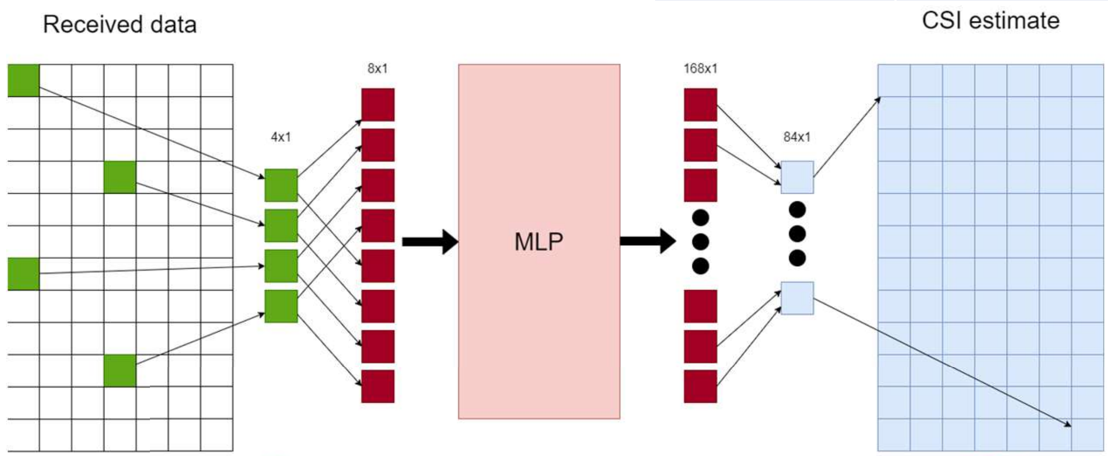


Dataset

SATNAC 2022

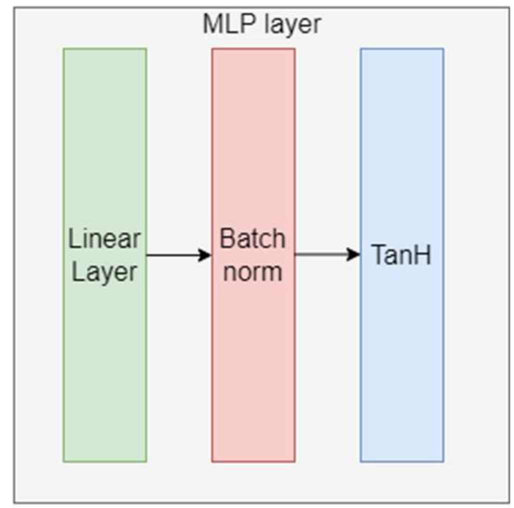
- All datasets and results are in a SISO antenna setup
- MLP input data is generated from the received pilot data
- Multiple datasets are used in this study
- LS and LTE-MMSE is used as comparisons

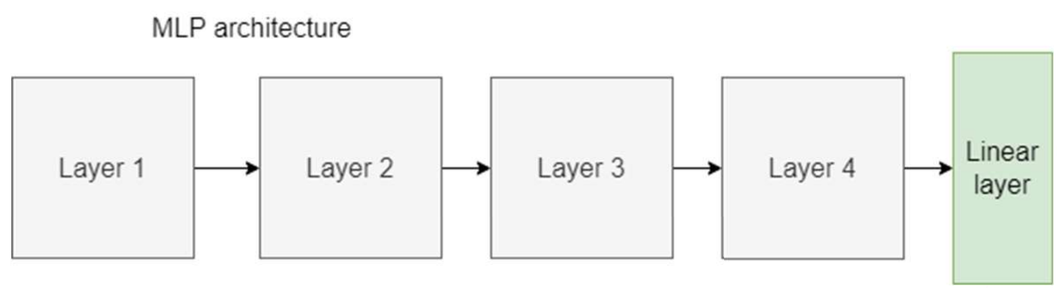
Dataset	Description
Noiseless data	Simple multi-path model
Noisy data	Simple multi-path model with 20dB AWGN receiver noise
EPA, EVA, ETU	Noisy data with delay profiles
Doppler fading	EVA dataset with Doppler fading



Architecture and training protocol

- Hyperparameters adjusted in this study:
  - Learning rate
  - Batch size
  - Network depth
  - Network width
  - Batch normalisation
  - Initialisation seeds
- Network are selected based on validation set loss and tested on test set BER.





Hosted by Telkom

Feature representation

- In our feature representation section we explore two features:
  - Angle and absolute
  - Quadrature and In-phase

Dataset	Features used	BER	Loss
Noiseless	Angle	8.33e-4	4.61e-2
Noisy (Trained on Noiseless)	Angle (no batch norm)	5.16e-3	2.31e-1
Noisy	Angle (no batch norm)	1.72e-2	1.28e-1
Noisy	QI	δ	5.23e-3

Feature representation

- We discover that the angle representation does not outperform LS as expected
- Applying the QI representation, which is Euclidian in nature, performs more predictably

Angle of the predicted CSI

-2,7734	-2,7734	-2,7734	-2,7734	-2,7734	-2,7734	-2,7734
-2,8164	-2,8164	-2,8164	-2,8164	-2,8164	-2,8164	-2,8164
-2,8593	-2,8593	-2,8593	-2,8593	-2,8593	-2,8593	-2,8593
-2,9023	-2,9023	-2,9023	-2,9023	-2,9023	-2,9023	-2,9023
-2,9452	-2,9452	-2,9452	-2,9452	-2,9452	-2,9452	-2,9452
-2,9882	-2,9882	-2,9882	-2,9882	-2,9882	-2,9882	-2,9882
-3,0311	-3,0311	-3,0311	-3,0311	-3,0311	-3,0311	-3,0311
-3,0741	-3,0741	-3,0741	-3,0741	-3,0741	-3,0741	-3,0741
-3,117	-3,117	-3,117	-3,117	-3,117	-3,117	-3,117
3,1232	3,1232	3,1232	3,1232	3,1232	3,1232	3,1232
3,0802	3,0802	3,0802	3,0802	3,0802	3,0802	3,0802
3,0373	3,0373	3,0373	3,0373	3,0373	3,0373	3,0373

Effect of hyperparameter choices: Network size

- Through our hyperparameter sweep we find that:
  - Larger networks perform better
  - Simply adding more layers decreases performance
  - Adding width to this problem helps the performance
- Thus, only a combination of width and depth provides optimal results

BEST PERFORMING STRUCTURAL HYPERPARAMETERS TEST LOSS RESULTS FOR MLP TRAINED ON NOISELESS DATA AND NO BATCH NORMALISATION LAYERS

	Depth			
Width	1	2	3	4
10	8.41e-2	9.45e-2	1.27e-1	1.95e-1
100	5.41e-2	4.46e-2	2.54e-2	1.14e-3
1 000	4.71e-2	1.65e-3	9.49e-8	6.33e-8

Effect of hyperparameter choices: Batch normalisation

- When removing batch normalisation from the angle representation network, the MLP performs as expected on the noiseless dataset.
- For this reason, we inspect the results of the QI representation with and without batch normalisation.

Dataset	Features used	Result (BER)
Noiseless	Angle	8.33e-4
Noiseless	Angle (no batch norm)	δ
Noisy	Angle (no batch norm)	1.72e-2
Noisy	QI	δ

TEST LOSS RESULTS OVER GROUPED TRAINING PARAMETER FOR NOISY DATA WITH QI VALUES AS FEATURES

	Learning rate			
Batch size	1e-5	1e-4	1e-3	1e-2
32	5.28e-3	5.22e-3	5.49e-3	2.64e-2
128	5.32e-3	5.22e-3	5.49e-3	1.40e-2
256	5.33e-3	5.26e-3	5.56e-3	9.37e-3
512	5.36e-3	5.31e-3	5.56e-3	7.59e-3

THE SAME SETUP AS IN TABLE III, EXCEPT WITH BATCH NORMALISATION

	Learning rate			
Batch size	1e-5	1e-4	1e-3	1e-2
32	5.47e-3	5.64e-3	5.37e-3	6.33e-3
128	5.47e-3	5.5e-3	5.36e-3	5.75e-3
256	5.48e-3	5.48e-3	5.47e-3	5.66e-3
512	5.46e-3	5.47e-3	5.57e-3	5.59e-3

Observed performance

- Implementing the four layers and 1 000 wide networks, using QI representation and batch normalisation, we test the following:
  - Generalisation between different delay profiles
  - The performance on more complex datasets

BER OF MLP NETWORKS TRAINED ON THE DELAY PROFILE DATASETS AND APPLIED TO THE SAME OR OTHER DELAY PROFILE TEST SETS

Trained on	EPA	EVA	ETU
EPA	3.97e-3	1.10e-2	3.63e-2
EVA	4.82e-3	6.88e-3	1.57e-2
ETU	6.69e-3	8.35e-3	1.16e-2
LS Method	1.20e-2	1.43e-2	3.28e-2

BER OF DIFFERENT EQUALISATION METHODS ON THE DOPPLER DATASET WITH VARIOUS AMOUNTS OF DOPPLER SHIFT

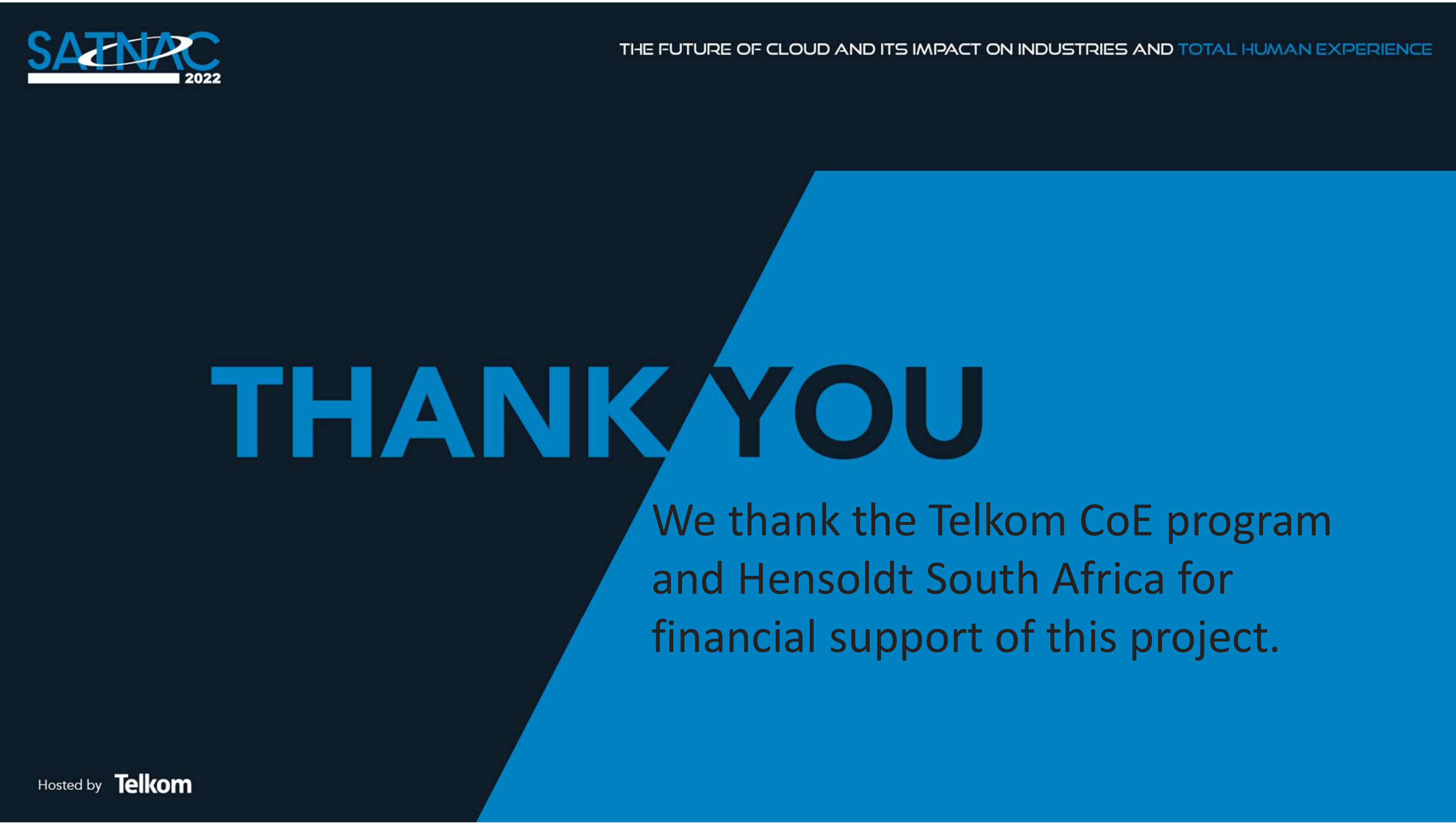
Doppler (Max Hz)	MLP	LS	MMSE
50	6.99e-3	1.38e-2	8.6e-3
100	8.01e-3	1.45e-2	8.6e-3
200	9.13e-3	1.5e-2	8.6e-3

Performance outcomes:

- **We observe some generalisation between different fading scenarios**
- **The MLPs compete with MMSE under Doppler fading conditions**

Conclusion

- An MLP outperforming LS and capable of contending with LTE-MMSE is developed
- These results could only have been obtained by:
  - Careful hyperparameters sweep for both architectural and training parameters
  - Inspecting the effect of batch normalisation with certain representations in relation to dataset complexity
- Finally, this research shows the importance of using Euclidian data representations when implementing MLPs


SATNAC 2022
THE FUTURE OF CLOUD AND ITS IMPACT ON INDUSTRIES AND TOTAL HUMAN EXPERIENCE
THANK YOU
We thank the Telkom CoE program and Hensoldt South Africa for financial support of this project.
Hosted by Telkom