

- **Transformer model for language understanding**
- **Scaled dot product attention**
- **Multi-head attention**
- **Point wise feed forward network**
- **Encoder and decoder**
- **Create the Transformer**
- **Set hyperparameters**
- **Optimizer**
- **Training and checkpointing**
- **Fine-tuning a BERT model**
- **The data**
- **Mask and input type**
- **Put it all together**
- **The model**
- **Set up the optimizer**
- **Train the model**
- **Re-encoding a large dataset**
- **Create `tf.data.Dataset` for training and evaluation**
- **TFModels BERT on TFHub**
- 19. Low level model building**
 - **Restore the weights:**
 - **Optimizers and schedules**