1. **Transformer model for language understanding**
2. **Scaled dot product attention**
3. **Multi-head attention**
4. **Point wise feed forward network**
5. **Encoder and decoder**
6. **Create the Transformer**
7. **Set hyperparameters**
8. **Optimizer**
9. **Training and checkpointing**
10. **Fine-tuning a BERT model**
11. **The data**
12. **Mask and input type**
13. **Put it all together**
12. **The model**
13. **Set up the optimizer**
14. **Train the model**
15. **Save the model**
16. **Re-encoding a large dataset**
17. **Create tf.data.Dataset for training and evaluation**
18. **TFModels BERT on TFHub**
19. **Low level model building**
20. **Restore the weights:**
21. **Test run it.**
22. **Optimizers and schedules**