




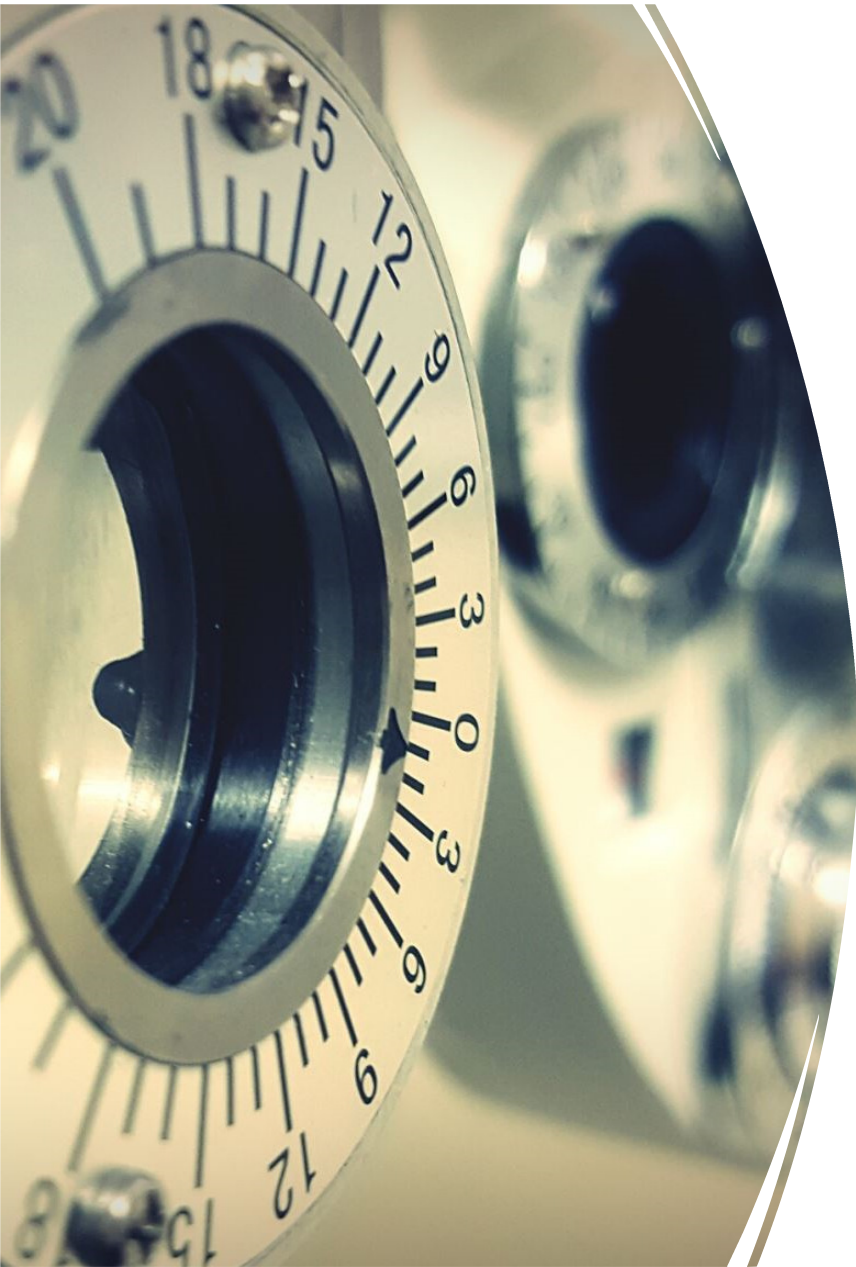
Transformers

- **Base “Model”:** Attention Is All You Need



NLP (Natural Language Processing)

- BERT (Bidirectional Encoder Representations from Transformers):
- GPT (Generative Pre-trained Transformer)
- T5 (Text-To-Text Transfer Transformer)
- XLNet
- RoBERTa
- ALBERT (A Lite BERT)
- DistilBERT
- ELECTRA





Computer Vision

- ViT (Vision Transformer)
- DETR (DEtection Transformer)


More

- CLIP (Contrastive Language–Image Pretraining)
- DALL-E
- MUM (Multitask Unified Model)
- Switch Transformer





Some
models we
are going to
use...



- `gpt2`
 - `bert-base-uncased`
 - `distilbert-base-cased-distilled-squad`
 - `facebook/bart-large-cnn`
- 