Deep Learning

Components

- Training data: (**x,** y) y: discrete quantity, real number
- Model:
    - Architecture
        - (FFNN) Input layer → (Hidden Layer)+ → Output layer
        - (CNNs)
        - (RNNs) GRU, LSTM
- Loss function
    - Regression: MSE, MAE
    - Binary classification: BCE
- Optimization procedure
    - SGD
- Model selection/evaluation
    - HPT
    - Regression: MSE
    - Classification: Precision, recall, F1-score, accuracy

Visual neural network

- https://playground.tensorflow.org

NLP, Images, Vector data

- Vector data → ML model → Label
- NLP → **Feature engineering** → ML model → Label
- Image → **Feature engineering** → ML model → Label

Text/Image/Speech → **Deep learning model** → Output Layer → Label / Sequence of tokens

- Sequence-to-sequence models

Recurrent neural networks (RNNs)

- **Text** → Sequence of tokens (words, characters, subword units)
- Sequence models
    - Hidden markov models (HMMs) {Generative models}
    - MaxEnt
    - Conditional random fields (CRFs) {Conditional models}
    - RNNs → GRUs, LSTM
    - Attention models
    - Transformers

Input:

- **Sequence** of tokens
- Feature vector

s_1: T<1>^{(1)} T<2>^{(1)} T<3>^{(1)} T<4>^{(1)} … T<k>^{(1)} (Review #1) | y^{(1)}
x_1 **x**_2 **x**_3 **x**_4 … **x**_k | y^{(1)}

s_2: T<1>^{(2)}, T<2>^{(2)} … T<k>^{(2)} (Review #2) | y^{(2)}
.
.
s_n: T<n>^{(2)}, T<n>^{(2)} … T<n>^{(2)} (Review #n) | y^{(n)}