Data Science Course

Lecture 26

Regularization

- Small values for parameters theta:
    - Simpler hypothesis
    - Less prone to overfitting.
- We take the cost function and we modify it to shrink all the parameters.

Model ensembling

- Consists of pooling together the predictions of a set of different models, to produce better predictions.
- Assumption: different good models trained independently are likely to be good for different reasons.
- Weighted average.
- Diversity.
- As good as possible and as different as possible.