Week 3

How to Win a Data Science Competition: Learn from Top Kagglers

Week 3

Discuss this week's modules here.

136 threads · Last post 6 days ago

Go to forum

Metrics Optimization

This week we will first study another component of the competitions: the evaluation metrics. We will recap the most prominent ones and then see, how we can efficiently optimize a metric given in a competition.

Key Concepts

- Describe the role of correct metric optimization method in a competition
- Analyze new metrics
- Create constant baselines
- Recall the most important classification and regression metrics
- Describe what libraries can be used to optimize a particular metric

Less

Metrics optimization

Reading: Week 3 overview 10 min

Video: Motivation 8 min — Resume

Video: Regression metrics review I 14 min

Notebook: Constants for MSE and MAE 10 min

Video: Regression metrics review II 8 min

Notebook: A note about weighted median 10 min

Video: Classification metrics review 20 min

Video: General approaches for metrics optimization 6 min

Video: Regression metrics optimization 10 min

Video: Classification metrics optimization I 7 min

Video: Classification metrics optimization II 6 min

Notebook: "Soft kappa" loss implementation 5 min

Practice Quiz: Metrics 6 questions

Quiz: Metrics 6 questions — Due Oct 12, 1:59 AM CDT

Reading: Comments on quiz 10 min

Reading: Additional material and links 10 min

Advanced Feature Engineering I

In this module we will study a very powerful technique for feature generation. It has a lot of names, but here we call it "mean encodings". We will see the intuition behind them, how to construct them, regularize and extend them.

Key Concepts

- Regularize mean encodings
- Extend mean encodings
- Summarize the concept of mean encodings

Less

Mean encodings

- **Video:** Concept of mean encoding 8 min — Resume
- **Video:** Regularization 7 min
- **Video:** Extensions and generalizations 10 min
- **Quiz:** Mean encodings 4 questions — Due Oct 12, 1:59 AM CDT
- **Reading:** Comments on quiz 10 min
- **Notebook:** Mean encoding implementations notebook
- **Programming Assignment:** Mean encoding implementation 3h — Due Oct 12, 1:59 AM CDT
- **Reading:** Final project advice #3 10 min