

Messages

Resources

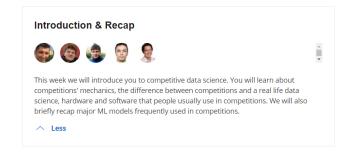


Yunpeng Cheng ~

### Week 1

What do you want to learn?

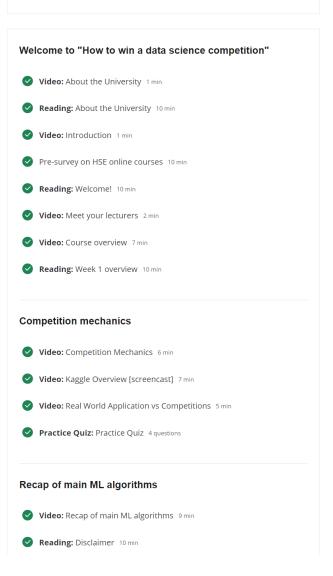
How to Win a Data Science Competition: Learn from Top Kagglers

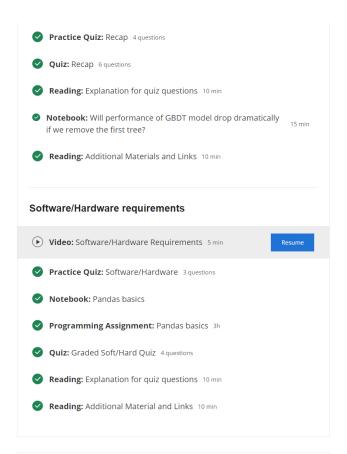


#### **Key Concepts**

- Describe competition mechanics
- Compare real life applications and competitions
- Summarize reasons to participate in data science competitions
- Describe main types of ML algorithms
- Describe typical hardware and software requirements
- Analyze decision boundaries of different classifiers
- Use standard ML libraries







# Feature Preprocessing and Generation with Respect to Models













In this module we will summarize approaches to work with features: preprocessing, generation and extraction. We will see, that the choice of the machine learning model impacts both preprocessing we apply to the features and our approach to generation of new ones. We will also discuss feature extraction from text with Bag Of Words and Word2vec, and feature extraction from images with Convolution Neural Networks.

^ Less

#### **Key Concepts**

- Explain how employed model impacts choice of preprocessing
- Summarize feature preprocessings for numeric and categorical features
- Summarize feature generation approaches for datetime and coordinates
- Summarize approaches to deal with missing values
- Outline the pipeline of applying Bag of Words
- Compare Bag of Words and Word2vec
- Explain how to extract CNN descriptors from images

∧ Less

## Feature preprocessing and generation with respect to models

▶ Video: Overview 6 min

Resum

▶ Video: Numeric features 13 min

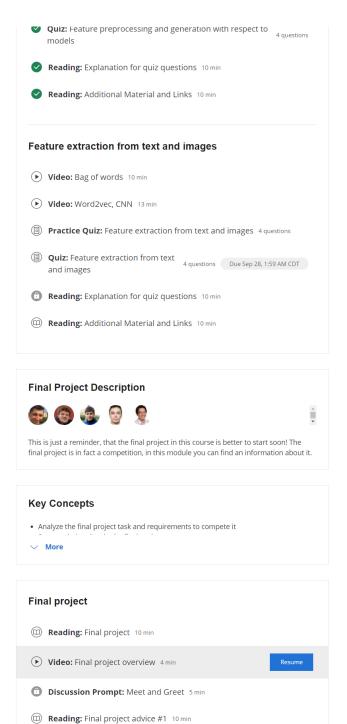
✓ Video: Categorical and ordinal features 10 min

**Video:** Datetime and coordinates 8 min

**Video:** Handling missing values 10 min

 Practice Quiz: Feature preprocessing and generation with respect to models

4 question



Survey

Survey 10 min