



Contribution ID: 24

Type: **not specified**

Coffee Seminar: Becoming a Data Scientist

Thursday, 20 September 2018 15:35 (55 minutes)

In this talk we'll cover materials and knowledge base of how to become a world class Data Scientists. We'll use real kaggle competition dataset and build for it a set of ML models. During this session we'll introduce advanced concepts of ML training, like feature embedding, data transformation and ML model fine-tuning. We'll discuss how to work large datasets, techniques to avoid RAM limitations problems, and how to create an ensemble of multiple ML models. We'll use XGBoost library and Keras ML framework to build our models and embeddings matrix. Even though the materials of this session is quite advance we expect students to listen and follow up with introduced concepts.

Primary author: Dr KUZNETSOV, Valentin (Cornell University (US))

Presenter: Dr KUZNETSOV, Valentin (Cornell University (US))

Session Classification: Coffee Seminar