## **Competition link:**

<https://www.kaggle.com/c/epamdscourseclassification3>

## **Methods**

### **LDA and QDA**

* Linear and Quadratic Discriminant Analysis (sklearn):
  + [base](https://scikit-learn.org/stable/modules/lda_qda.html)
* Linear discriminant analysis:
  + [further reading](https://en.wikipedia.org/wiki/Linear_discriminant_analysis)
* The Elements of Statistical Learning, **p. 106-119**:
  + [advanced - “The Elements of Statistical Learning”, Hastie T., Tibshirani R., Friedman J., Section 4.3, p.106-119](https://web.stanford.edu/~hastie/ElemStatLearn/printings/ESLII_print12.pdf)

### **Gradient boosting**

* A Gentle Introduction to Gradient Boosting, Cheng Li
  + [base](http://www.chengli.io/tutorials/gradient_boosting.pdf)
* Boosting article by ODS:
  + <https://mlcourse.ai/articles/topic10-boosting/>
* Gradient Boosting (Wikipedia)
  + [further reading](https://en.wikipedia.org/wiki/Gradient_boosting)
* Gradient Boosting explained [demonstration]
  + [visualization](http://arogozhnikov.github.io/2016/06/24/gradient_boosting_explained.html)
* The Elements of Statistical Learning, **p. 337**:
  + [advanced - “The Elements of Statistical Learning”, Hastie T., Tibshirani R., Friedman J., Section 10, p.337](https://web.stanford.edu/~hastie/ElemStatLearn/printings/ESLII_print12.pdf)

### **Random Forest**

* An Implementation and Explanation of the Random Forest in Python
  + [base](https://towardsdatascience.com/an-implementation-and-explanation-of-the-random-forest-in-python-77bf308a9b76)
* The Elements of Statistical Learning, Section 15, **p.587**
  + [advanced](https://web.stanford.edu/~hastie/ElemStatLearn/printings/ESLII_print12.pdf)

## **Ensembling methods**

* Comprehensive Guide to Ensemble Learning
  + [Overview](https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-for-ensemble-models/)

## **Class imbalance**

* About metrics
  + [1](https://towardsdatascience.com/what-metrics-should-we-use-on-imbalanced-data-set-precision-recall-roc-e2e79252aeba)
  + [2](https://machinelearningmastery.com/tour-of-evaluation-metrics-for-imbalanced-classification/)
* Common strategies to resolve the problem
  + [strategies](https://machinelearningmastery.com/tactics-to-combat-imbalanced-classes-in-your-machine-learning-dataset/)

## **Classification problems**

* Multiclass and multilabel
  + [Overview](https://scikit-learn.org/stable/modules/multiclass.html)
* Metrics
  + [read part2 only](https://towardsdatascience.com/journey-to-the-center-of-multi-label-classification-384c40229bff)