

캐글 코리아
캐글 스터디
커널 커리큘럼

작성자 프로필

카이스트
생명화학공학과
분자 시뮬레이션 실험실
(Prof. Jihan Kim)

Ph.D candidate 이유한

<https://www.kaggle.com/youhanlee>

<https://www.facebook.com/groups/KaggleKoreaOpenGroup/>

이 커리큘럼 참여 방법

필사적으로 필사하세요

**커널의 A 부터 Z 까지 다 똑같이
따라 적기!**

**똑같이 3번적이고 다음 커널로
넘어가시면 됩니다.**

Binary classification

- Tabular data -

1st level. Titanic competition

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



1st kernel

<http://kaggle-kr.tistory.com/category/Kaggle%20ED%8A%9CED%86%A0%EB%A6%AC%EC%96%BC/Titanic%20ED%8A%9CED%86%A0%EB%A6%AC%EC%96%BC>

2nd kernel

<https://www.kaggle.com/ash316/eda-to-prediction-dietanic>

3rd kernel

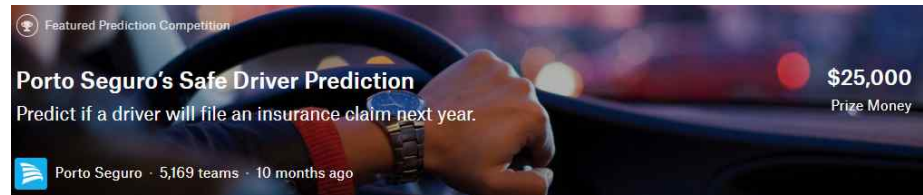
<https://www.kaggle.com/yassineghouzam/titanic-top-4-with-ensemble-modeling>

4th kernel

<https://www.kaggle.com/arthurtok/introduction-to-ensembling-stacking-in-python>

2nd level. Porto

<https://www.kaggle.com/c/porto-seguro-safe-driver-prediction>



1st kernel
<https://www.kaggle.com/bertcarremans/data-preparation-exploration>

2nd kernel
<https://www.kaggle.com/arthurtok/interactive-porto-insights-a-plot-ly-tutorial>

3rd kernel
<https://www.kaggle.com/aharless/xgboost-cv-lb-284>

4th kernel
<https://www.kaggle.com/gpreda/porto-seguro-exploratory-analysis-and-prediction>

이 외에 상위권
솔루션들이 많으니
따로 더 공부하셔도
됩니다!

3rd level. Home credit competition

<https://www.kaggle.com/c/home-credit-default-risk>



1st kernel
<https://www.kaggle.com/willkoehrsen/start-here-a-gentle-introduction>

2nd kernel
<https://www.kaggle.com/willkoehrsen/introduction-to-manual-feature-engineering>

3rd kernel
<https://www.kaggle.com/eliotbarr/stacking-test-sklearn-xgboost-catboost-lightgbm>

4th kernel
<https://www.kaggle.com/jsaguiar/lightgbm-7th-place-solution>

이 외에 상위권
솔루션들이 많으니
따로 더 공부하셔도
됩니다!

Multi-class classification

- Tabular data -

1st level. Costa-rican competition

<https://www.kaggle.com/c/costa-rican-household-poverty-prediction>



1st kernel
<https://www.kaggle.com/willkoehrsen/a-complete-introduction-and-walkthrough>

2nd kernel
<https://www.kaggle.com/youhanlee/3250feats-532-feats-using-shap-lb-0-436>

3rd kernel
<https://www.kaggle.com/skooch/xgboost>

Binary classification

- Image classification -

1st level. Statoil competition

<https://www.kaggle.com/c/statoil-iceberg-classifier-challenge>



1st kernel

<https://www.kaggle.com/devm2024/keras-model-for-beginners-0-210-on-lb-eda-r-d>

2nd kernel

<https://www.kaggle.com/devm2024/transfer-learning-with-vgg-16-cnn-aug-lb-0-1712>

3rd kernel

<https://www.kaggle.com/submarineering/submarineering-even-better-public-score-until-now>

4th kernel

<https://www.kaggle.com/wvadim/keras-tf-lb-0-18>

Multi-class classification

- Image classification -

1st level. Fruits 360

<https://www.kaggle.com/uciml/mushroom-classification>



1st kernel

<https://www.kaggle.com/amadeus1996/fruits-360-transfer-learning-using-keras>

2nd level. Fruits 360

<https://www.kaggle.com/zalando-research/fashionmnist>



1st kernel

<https://www.kaggle.com/shivamb/how-autoencoders-work-intro-and-usecases>

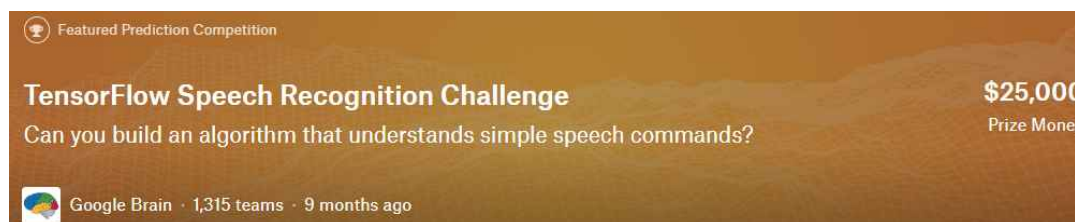
2nd kernel

<https://www.kaggle.com/bugraokcu/cnn-with-keras>

Audio recognition

1st level. Tensorflow competition

<https://www.kaggle.com/c/tensorflow-speech-recognition-challenge>



Featured Prediction Competition

TensorFlow Speech Recognition Challenge

Can you build an algorithm that understands simple speech commands?

\$25,000
Prize Money

Google Brain · 1,315 teams · 9 months ago

1st kernel
<https://www.kaggle.com/davids1992/speech-representation-and-data-exploration>

2nd kernel
<https://www.kaggle.com/alphasis/light-weight-cnn-lb-0-74>

3rd kernel
<https://www.kaggle.com/kcs93023/tf-speech-recognition-by-cnn>

4th kernel
<https://www.kaggle.com/ivallesp/wavception-v1-a-1-d-inception-approach-lb-0-76>

Regression - Tabular data -

1st level. New-York taxi competition

<https://www.kaggle.com/c/nyc-taxi-trip-duration>



1st kernel
<https://www.kaggle.com/gaborfodor/from-eda-to-the-top-lb-0-367>

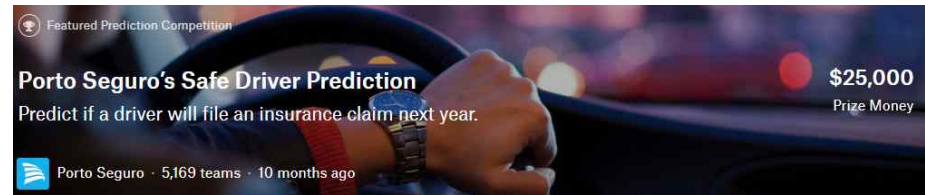
2nd kernel
<https://www.kaggle.com/drgilermo/dynamics-of-new-york-city-animation>

3rd kernel
<https://www.kaggle.com/aiswaryaramachandran/eda-baseline-model-0-40-rmse>

4th kernel
<https://www.kaggle.com/danijelk/beat-the-benchmark>

2nd level. Zillow competition

<https://www.kaggle.com/c/zillow-prize-1>



1st kernel

<https://www.kaggle.com/sudalairajkumar/simple-exploration-notebook-zillow-prize>

2nd kernel

<https://www.kaggle.com/anokas/simple-xgboost-starter-0-0655>

3rd kernel

<https://www.kaggle.com/viveksrinivasan/zillow-eda-on-missing-values-multicollinearity>

4th kernel

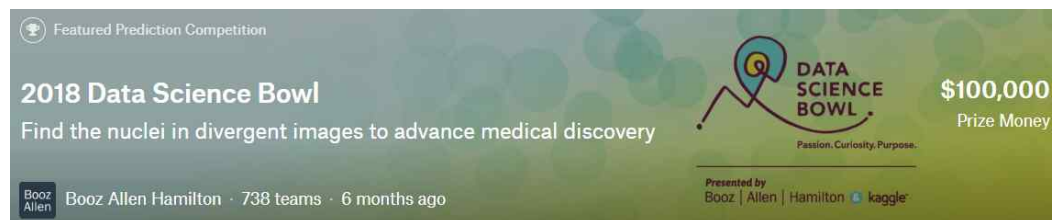
<https://www.kaggle.com/aharless/xgboost-lightgbm-and-ols-and-nn>

이 외에 상위권
솔루션들이 많으니
따로 더 공부하셔도
됩니다!

Object segmentation

1st level. Data bowl 2018

<https://www.kaggle.com/c/data-science-bowl-2018>



1st kernel
<https://www.kaggle.com/stkbailey/teaching-notebook-for-total-imaging-newbies>

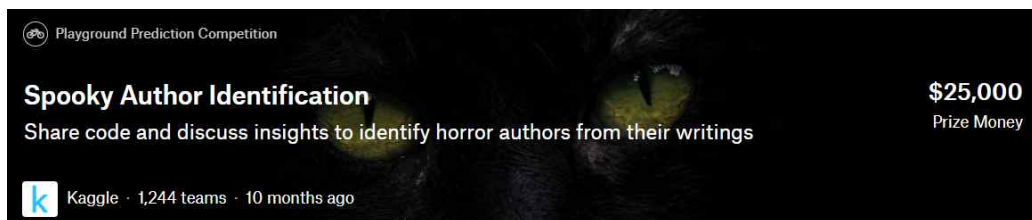
2nd kernel
<https://www.kaggle.com/keegil/keras-u-net-starter-lb-0-277>

3rd kernel
<https://www.kaggle.com/kmader/nuclei-overview-to-submission>

Natural language processing - classification, regression -

1st level. Spooky competition

<https://www.kaggle.com/c/spooky-author-identification>



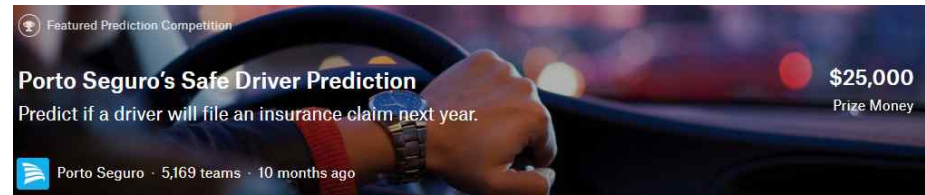
1st kernel
<https://www.kaggle.com/arthurtok/spooky-nlp-and-topic-modelling-tutorial>

2nd kernel
<https://www.kaggle.com/abhishek/approaching-almost-any-nlp-problem-on-kaggle>

3rd kernel
<https://www.kaggle.com/sudalairajkumar/simple-feature-engg-notebook-spooky-author>

2nd level. Mercari

<https://www.kaggle.com/c/mercari-price-suggestion-challenge>



1st kernel
<https://www.kaggle.com/thykhuely/mercari-interactive-eda-topic-modelling>

2nd kernel
<https://www.kaggle.com/knowledgegrappler/a-simple-nn-solution-with-keras-0-48611-pl>

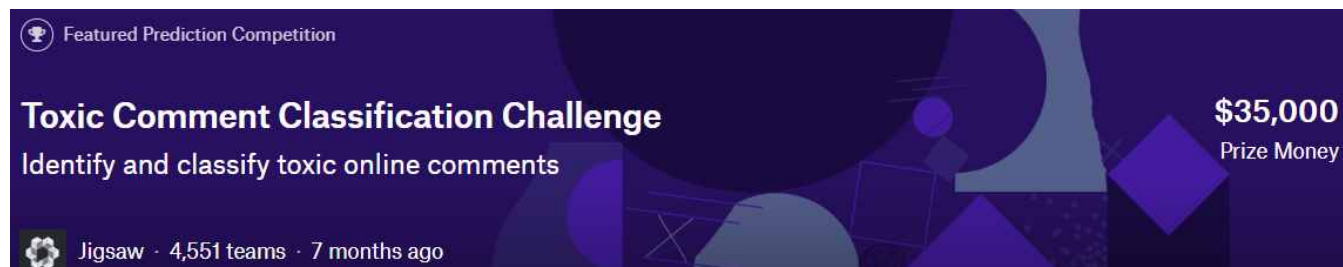
3rd kernel
<https://www.kaggle.com/rumbok/ridge-lb-0-41944>

4th kernel
<https://www.kaggle.com/peterhurford/lgb-and-fm-18th-place-0-40604>

이 외에 상위권
솔루션들이 많으니
따로 더 공부하셔도
됩니다!

3rd level. Toxic competition

<https://www.kaggle.com/c/jigsaw-toxic-comment-classification-challenge>



1st kernel
<https://www.kaggle.com/sbongo/for-beginners-tackling-toxic-using-keras>

2nd kernel
<https://www.kaggle.com/jagangupta/stop-the-s-toxic-comments-eda>

3rd kernel
<https://www.kaggle.com/tunguz/logistic-regression-with-words-and-char-n-grams>

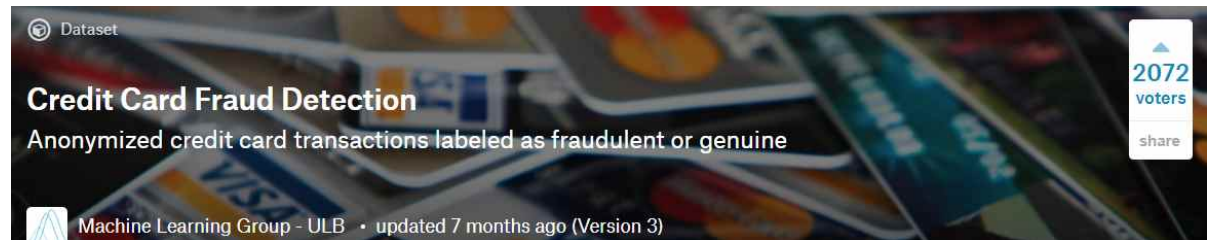
4th kernel
<https://www.kaggle.com/rhodiumbeng/classifying-multi-label-comments-0-9741-lb>

이 외에 상위권
솔루션들이 많으니
따로 더 공부하셔도
됩니다!

**Other dataset
- anomaly detection,
visualization-**

1st level. Credit card

<https://www.kaggle.com/mlg-ulb/creditcardfraud>



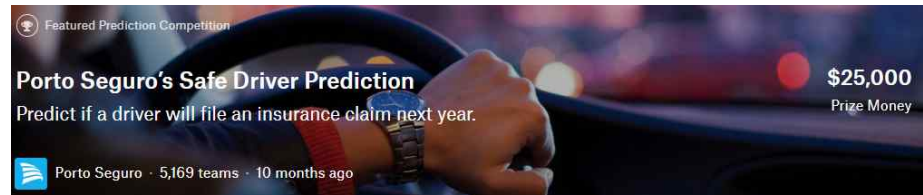
1st kernel
<https://www.kaggle.com/joparga3/in-depth-skewed-data-classif-93-recall-acc-now>

2nd kernel
<https://www.kaggle.com/pavansanagapati/anomaly-detection-credit-card-fraud-analysis>

3rd kernel
<https://www.kaggle.com/matheusfacure/semi-supervised-anomaly-detection-survey>

2nd level. Kaggle survey

<https://www.kaggle.com/kaggle/kaggle-survey-2017>



1st kernel
<https://www.kaggle.com/ash316/novice-to-grandmaster>

2nd kernel
<https://www.kaggle.com/mhjabri/what-do-kagglers-say-about-data-science>

3rd kernel
<https://www.kaggle.com/hakkisimsek/plotly-tutorial-1>