

Mechanisms of Action (MoA) Prediction

(Kaggle competition)

Romet Aidla, Kadri Onemar, Linda Katariina Greents, Marge Maidla
Team name on Kaggle: Drugdealers

15.12.2020

Goal of the competition

To advance drug development through improvements to MoA prediction algorithms*

Our team's goal: not the top place on leaderboard but *learn as much as possible*

* MoA - Mechanism of Action refers to the specific biochemical interaction through which a drug substance produces its pharmacological effect.



Source: https://attract-eu.com/wp-content/uploads/2020/06/ATTRACT_drugdevelopment.png

Details of the project

- Duration: 3 Sept - 30 Nov 2020
- Unique dataset combining gene expression and cell viability data
- Develop an algorithm that labels each case in the test set as one or more MoA classes -> **multi-label classification problem**
- Evaluation: log loss function
- Team with best score on private test set wins the competition

Approach and results

- Each team member develops a model, afterwards perform blending
 - Two Keras NN-s (0.01892 and 0.01911), XGBoost(0.01998), Logistic Regression (0.02688)
- Detailed results:

ensemble-weights
(version 14/14)
13 days ago by Romet

From Notebook [ensemble-weights]

1789

▲ 449

Drugdealers

In the money

Gold

Silver

Bronze

0.01650

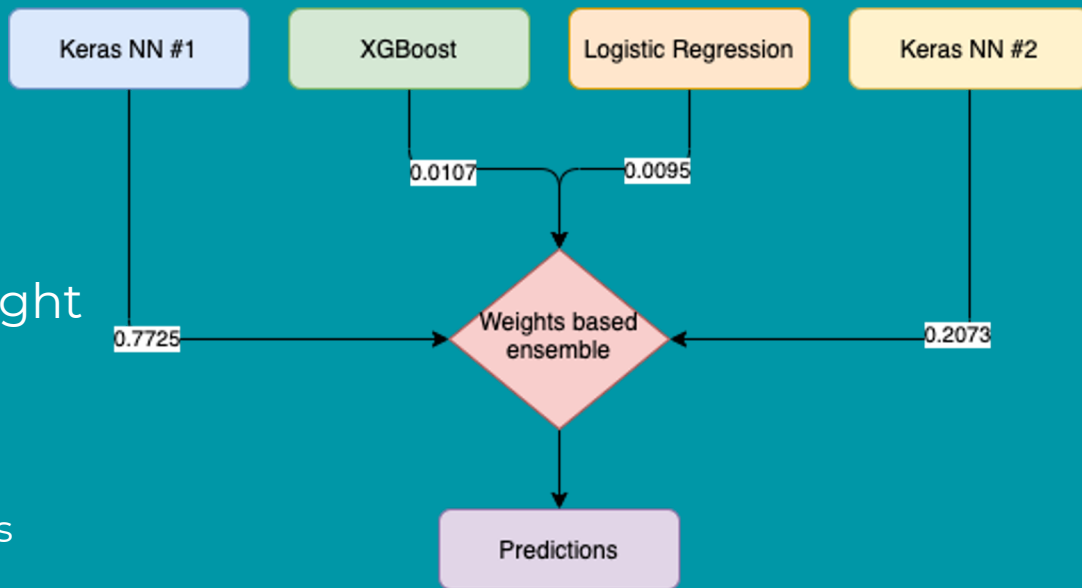
40

13d

#	Δpub	Team Name	Notebook	Team Members	Score	Entries	Last
1	▲ 6	Hungry for gold 🏆	◁ Fork of Blending w...	<div><div></div><div></div><div></div><div></div></div>	0.01599	216	13d
2	▲ 2	tmp		<div><div></div><div></div></div>	0.01599	136	13d
3	▲ 3	Duck Quake		<div><div></div><div></div></div>	0.01600	116	13d
4	▲ 8	Kanna Hashimoto with Fri...	◁ nn-svm-tabnet...	<div><div></div><div></div><div></div><div></div></div>	0.01600	50	13d

Ensemble

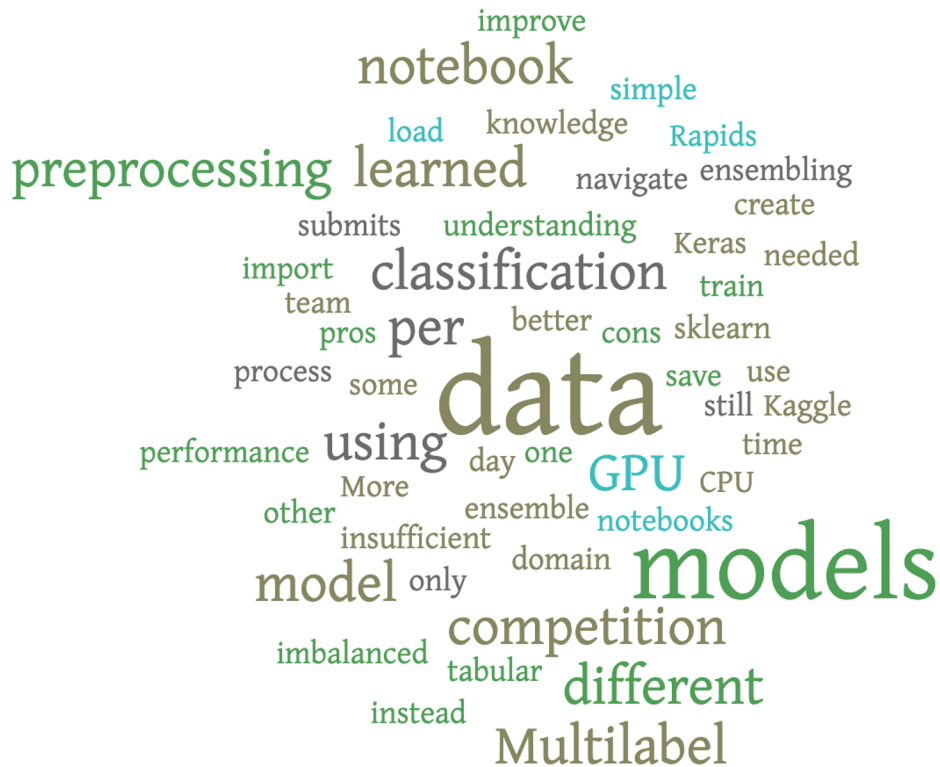
- Weights calculated by minimizing OOF losses
- Models not used (0 weight from optimizer):
 - Random Forest (Rapids)
 - KNN (Rapids)
 - Support Vector Machines (Rapids)
- Models not used (loss above baseline):
 - Catboost



Lessons learned

We all learned about:

- Kaggle competition process
- how to work with notebooks
- new models
- RAPIDS
- multilabel classification
- data preprocessing
- basics on drug development



References

Code repository:

- <https://github.com/rometaidla/mechanism-of-action>

Competition:

- <https://www.kaggle.com/c/lish-moa/overview>

Thank you!