Project report

MAGNETIC MATERIALS

Prediction of Curie temperature using Machine Learning

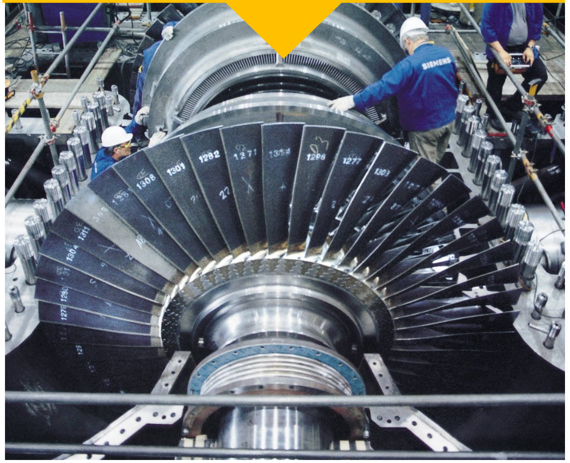
Outline

Plan for today

1. Project motivation

Industrial use of magnets

ENERGY SYSTEMS



MEDICINE

AUTOMOTIVE INDUSTRY

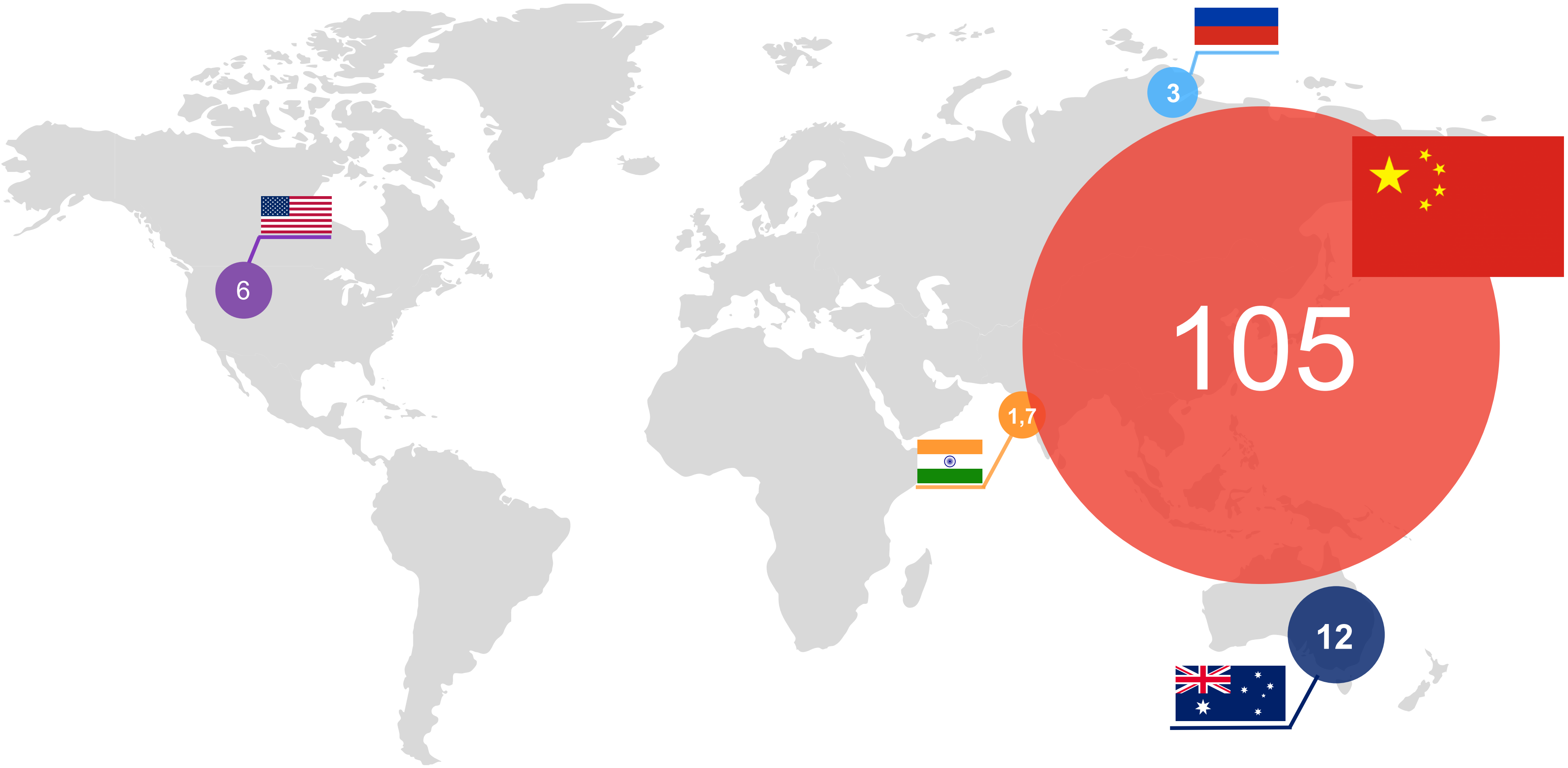
ELECTRONIC DEVICES

1. Critical materials

Rare Earth Elements mining.

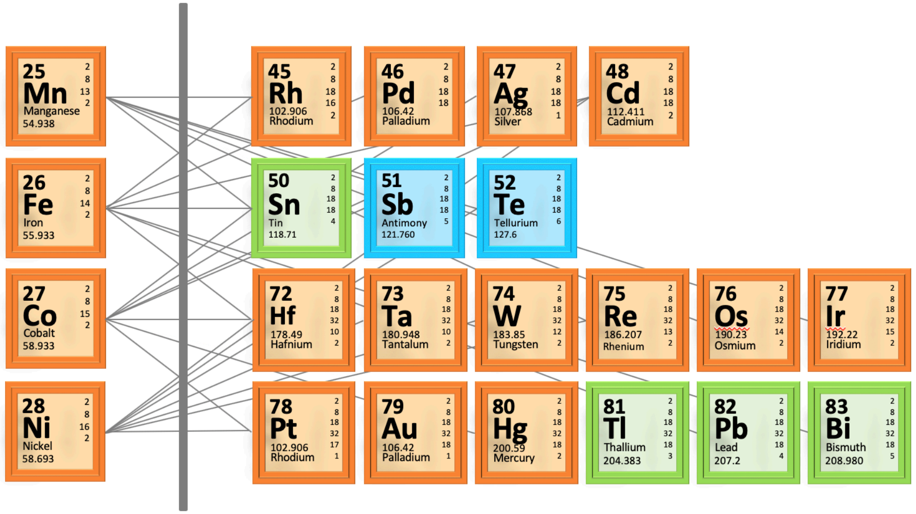
Production volume in thousand tons

About 90% of all rare earth elements are produced in China.



1. Computational search

Searching for a nowel magnetic materials



- Typically magnet might be considered as “useful”, only if its Curie temperature is more than 600K
- Curie temperature is one of the most important optimization parameters in a computational search for novel magnetic materials.
- Unfortunately, all existing modern methods can not exact information about Tc

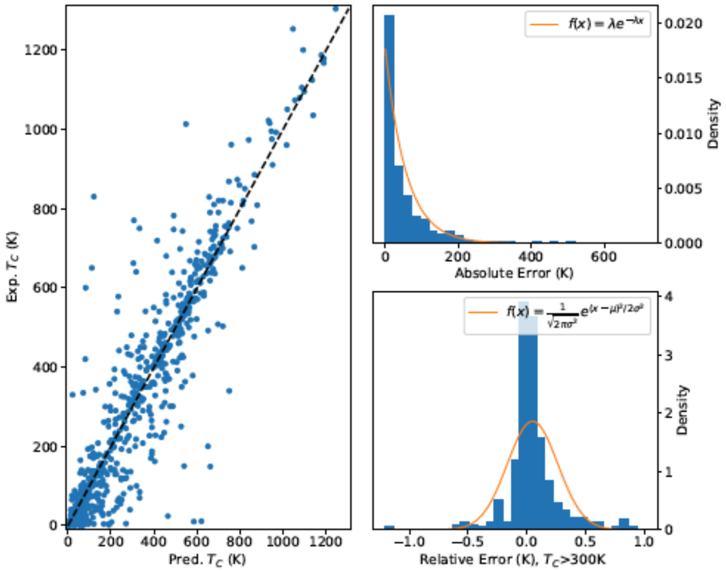
Computational prediction of new magnetic materials. Saeed Rahmanian Koshkaki, Zahed Allahyari, Artem R. Oganov 2020

2. Related works

Critical temperatures predictions using Machine Learning

Predicting the Curie temperature of ferromagnets using machine learning

James Nelson* and Stefano Sanvito†
School of Physics, AMBER and CRANN Institute, Trinity College, Dublin 2, Ireland
(Dated: June 21, 2019)



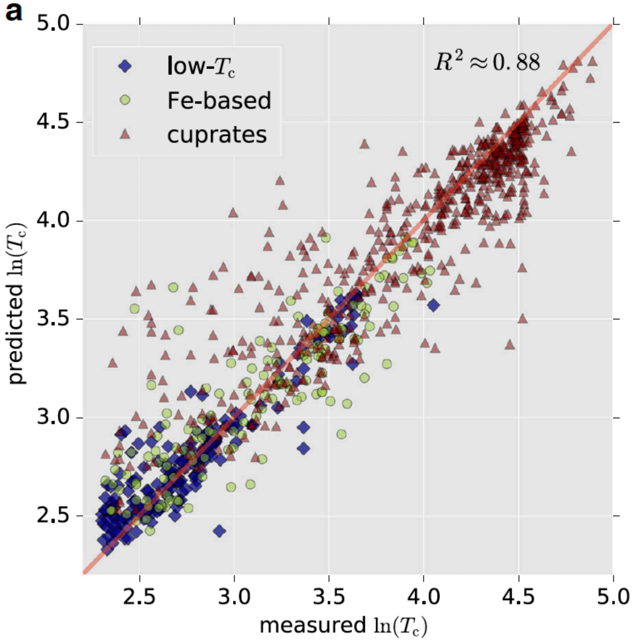
- Only Regresion problem
- Descriptors based on chemical formula
- Small dataset 2750 samples
- Best score 0.81
- RF model

ARTICLE OPEN

Machine learning modeling of superconducting critical temperature

Valentin Stanev[1,2], Corey Oses[3,4], A. Gilad Kusne[1,5], Efrain Rodriguez[2,6], Johnpierre Paglione[2,7], Stefano Curtarolo[3,4,8] and Ichiro Takeuchi[1,2]

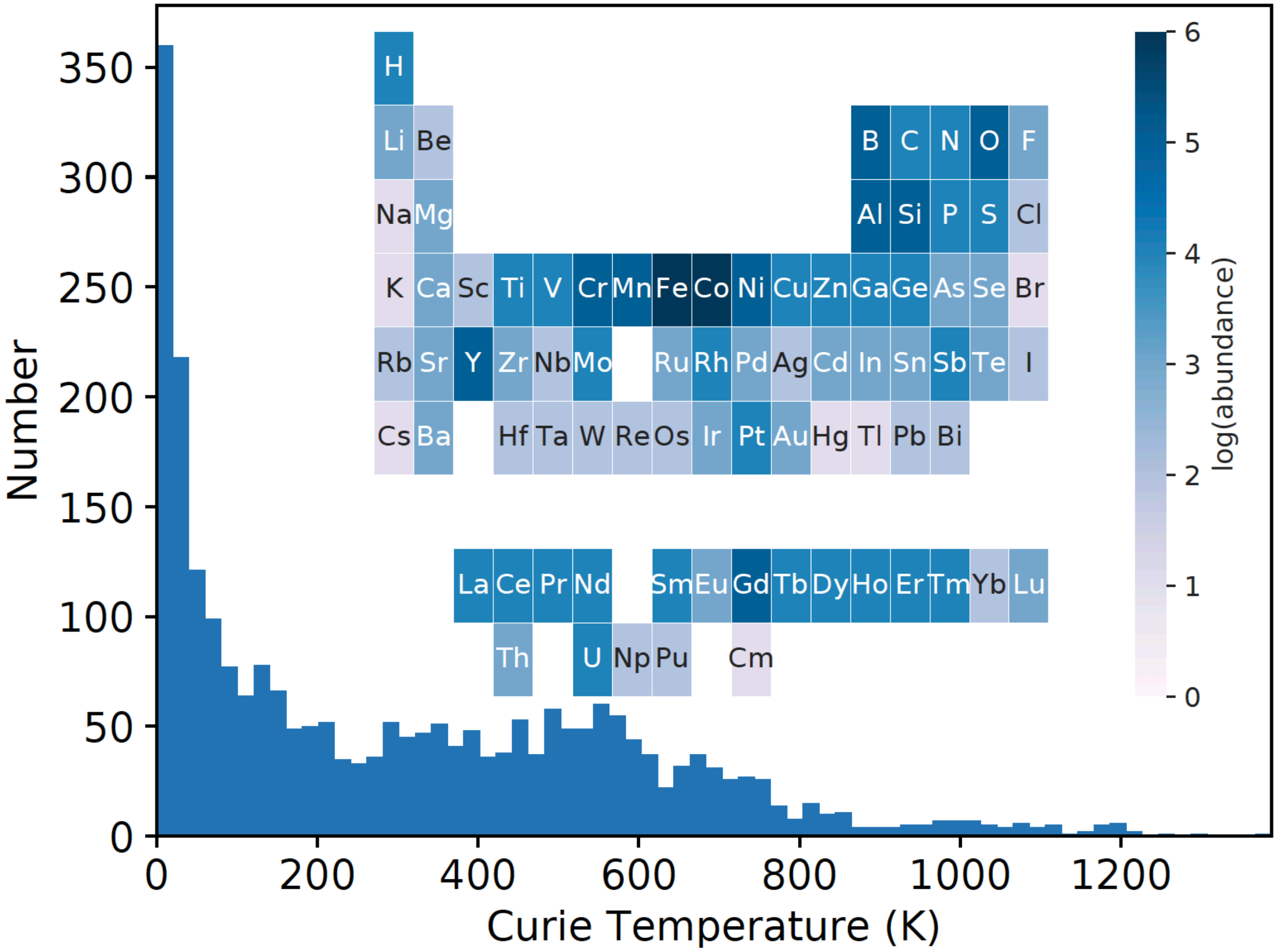
a



- Regresion and classification problems
- Descriptors based on chemical formula
- >12 000 Samples in Dataset
- Best R^2 score 0.88
- RF model

3.1 Dataset details

Critical magnetic temperature experimental data



5091 samples from experimental data

2557 samples after dropping of duplicates

Only chemical formula and Curie Temperature

	formula	T_c
0	(Ce0.05La0.95)3Al11	1.3
1	(Ce0.7La0.3)3Al11	3.7
2	(Ce0.9La0.1)3Al11	5.4
3	(Co0.38Fe0.62)2P	459.0
4	(Co0.4Fe0.6)2P	453.0

3.2 Dataset preprocessing

Data cleaning and generation of the descriptors

Attributes which might basically determined from the periodic law

1. *Atomic Weight*
2. *Number*
3. *Column*
4. *Row*
5. *Covalent Radius*
6. *Electronegativity*
7. *Bandgap*
8. *Magmom*
9. *MeltingT*
10. *Mendeleev Number*
11. *Space Group*
12. *Volume*

\+

Electronic structure attributes, which are the fraction of electrons from the s,p,d and f shells

1. *N Unfilled*
2. *N Valence*
3. *N d Unfilled*
4. *N d Valence*
5. *N f Unfilled*
6. *N f Valence*
7. *N p Unfilled*
8. *N p Valence*
9. *N s Unfilled*
10. *N s Valence*

×

For all descriptors from the first and second category applied statistics functions

1. *minimum*
2. *maximum*
3. *range*
4. *mode*
5. *mean*
6. *mean abs. dev.*

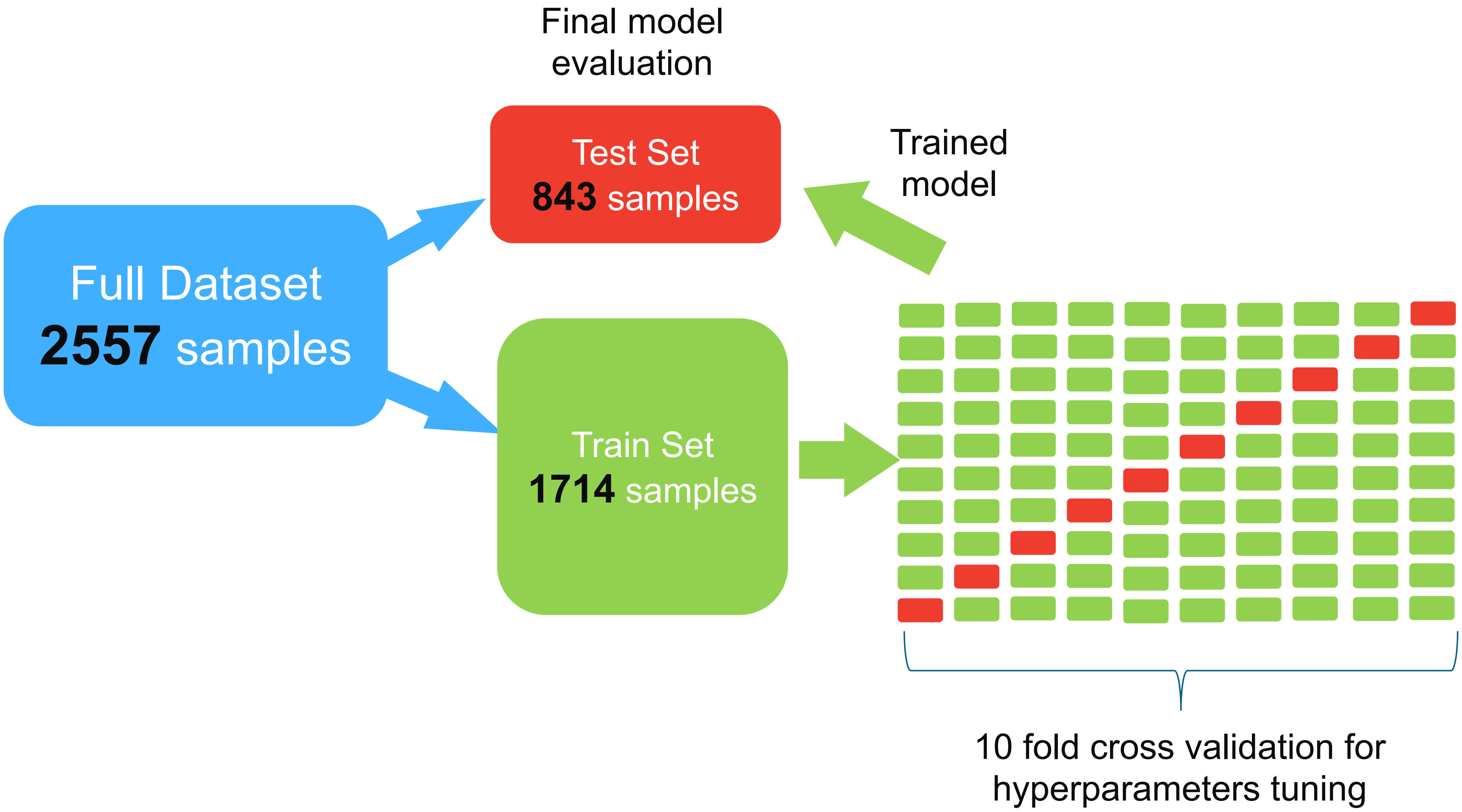
$+ V_{chem} = 215$

Total dimension of descriptors

[1] Ward, Logan et al. "A General-Purpose Machine Learning Framework for Predicting Properties of Inorganic Materials." Computational Materials 2.1 (2016)

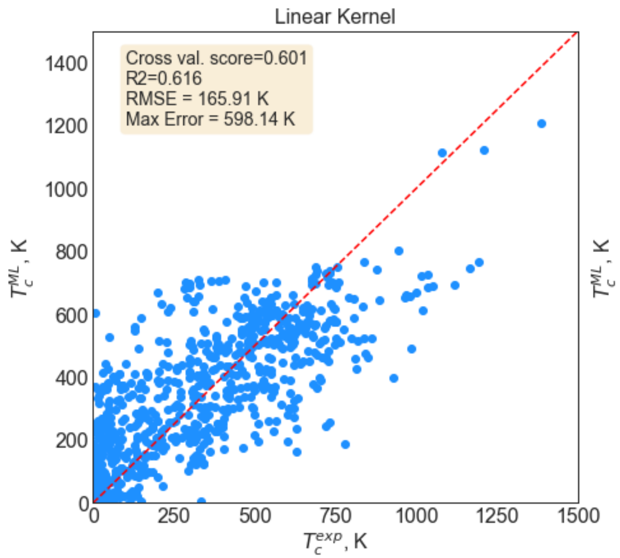
[2] L., Dunn, A., Faghaninia, A., Zimmermann, N. E. R., Bajaj, S, Matminer: An open source toolkit for materials data mining. Comput. Mater. Sci. 152, 60-69 (2018).

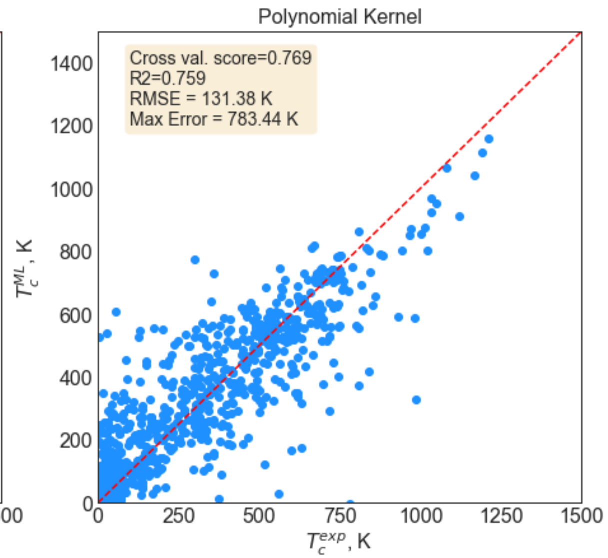
3.2 Models Validation

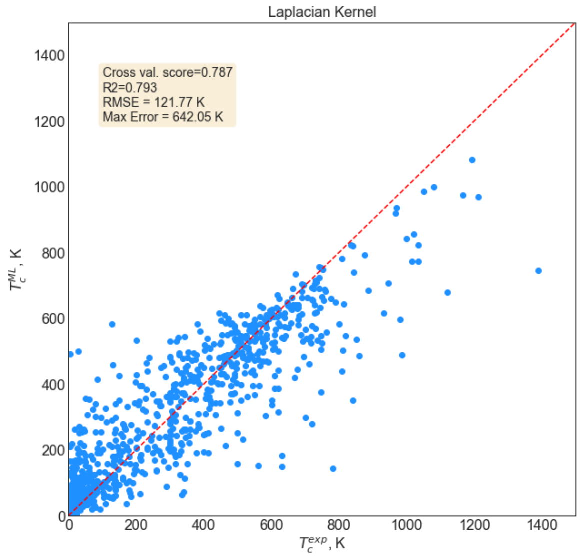


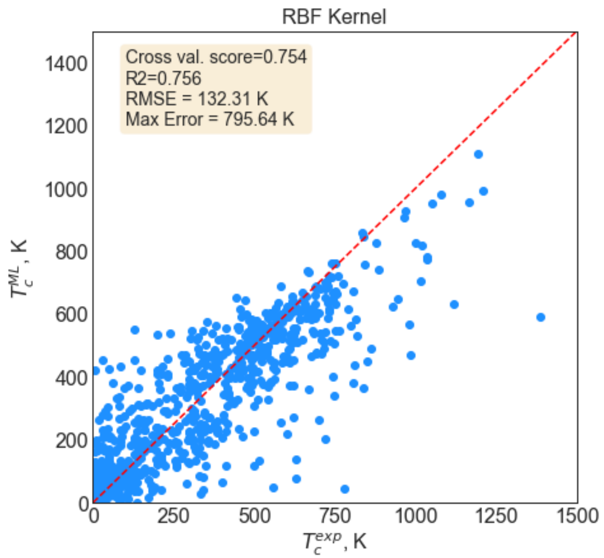
4.1 Kernel Ridge Regression

Model tuning and final perfomance







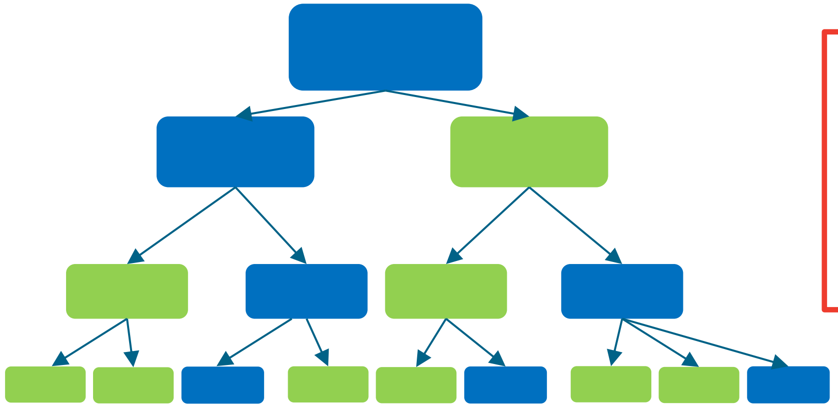


Metrics	Linear	Polynomial	RBF	Laplacian
R^2 (test)	0.616	0.759	0.756	0.793
R^2 (cross.val)	0.601	0.768	0.754	0.787
RMSE (K)	161.9	131.4	132.3	121.7
Max Error (K)	598.1	783.4	795.6	642.05

4.2 Random Forest Regressor

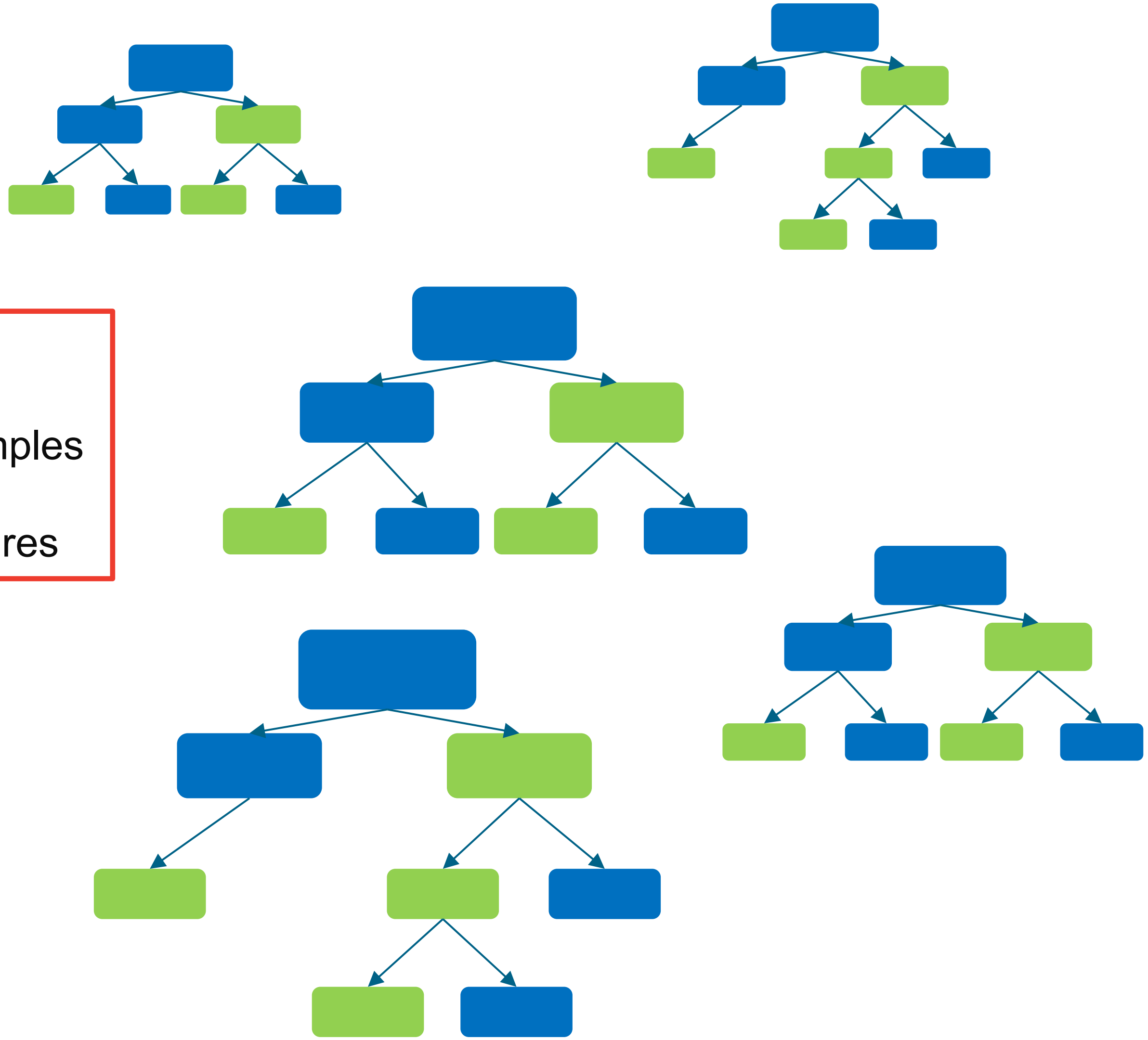
Basics of the alghoritm

Tree model



Bootsrapping
=
Random subset of samples
+
Random subset features

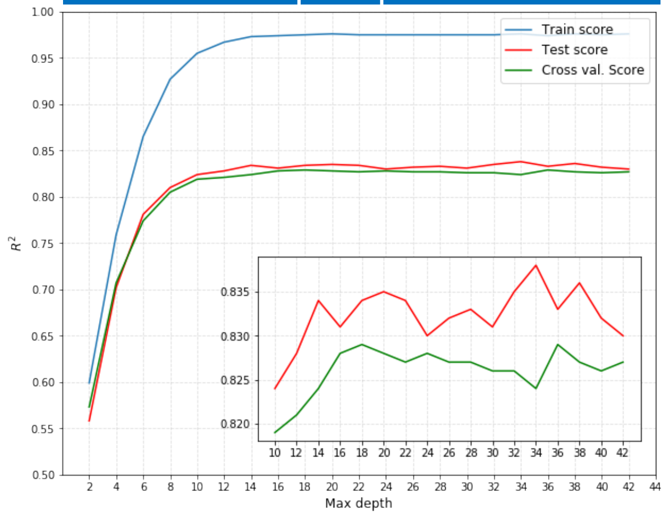
**Ensemble of trees =
Random Forest**



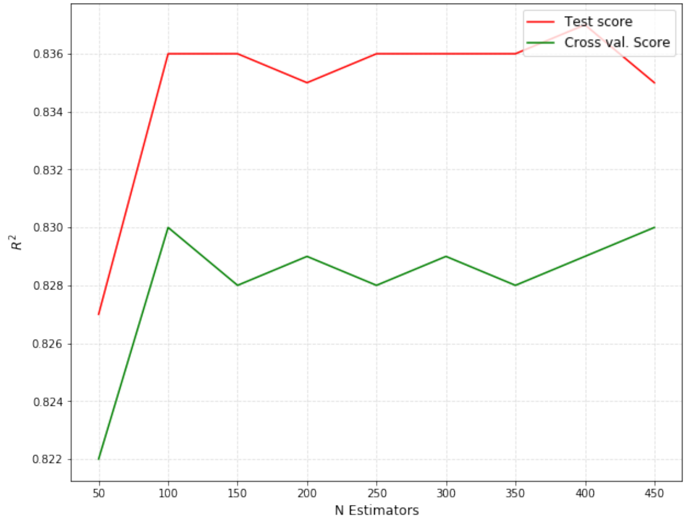
4.2 Random Forest Regressor

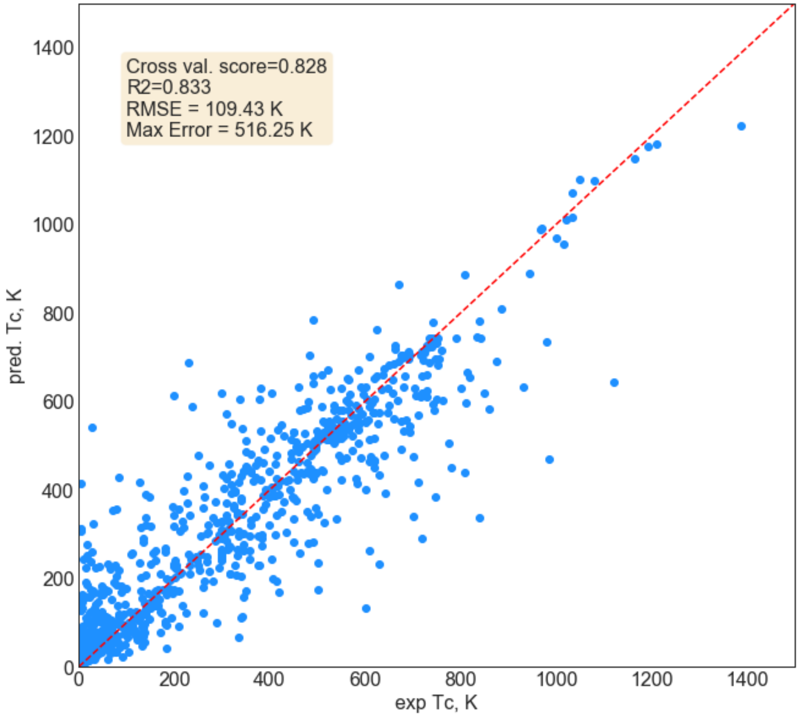
Hyperparameters tuning and model perfomance

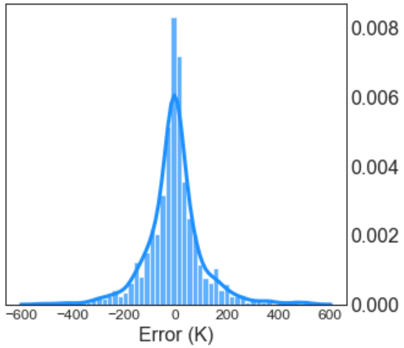
Tree depth optimization

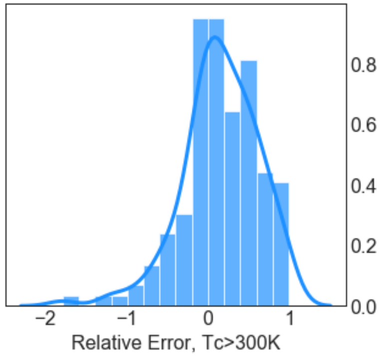


Number of Trees optimization





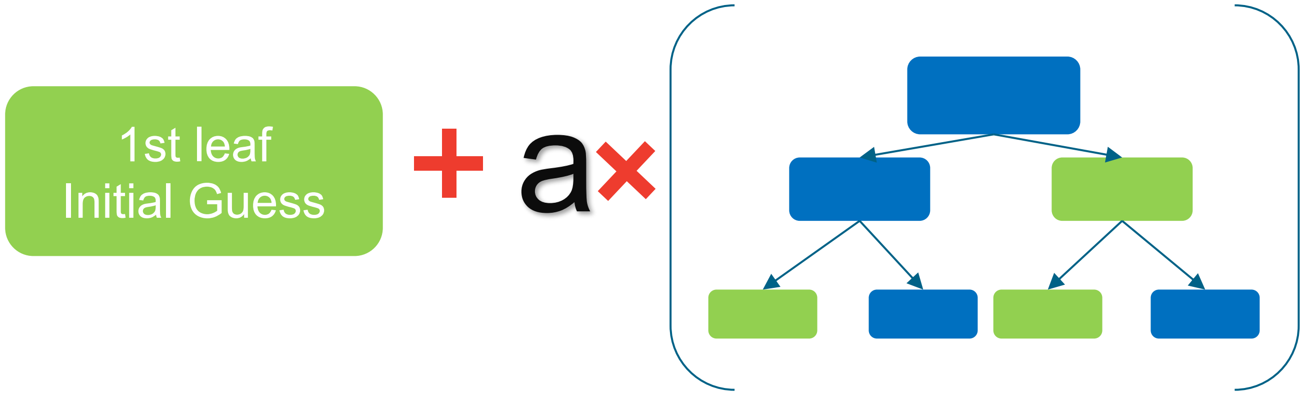


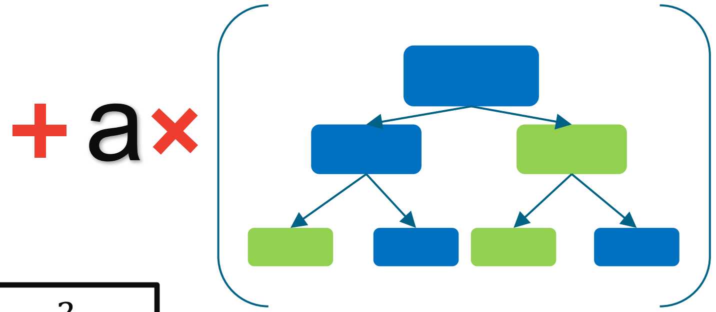


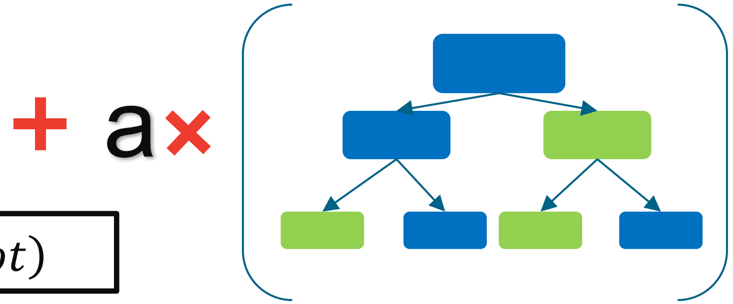
Metrics	All
R^2 (test)	0.837
R^2 (cross.val)	0.859
RMSE (K)	101.04
Max Error (K)	468.08

4.3 XGBoost

Basics of alghoritm







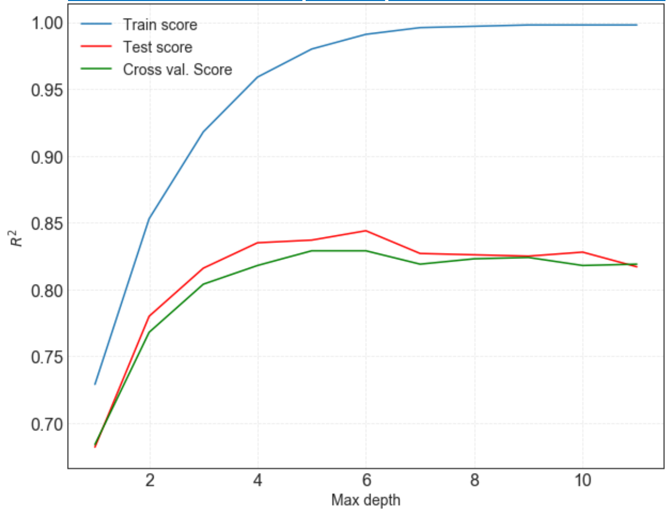
$$Similarity\ Score = \frac{(Sum\ of\ Residuals)^2}{Number\ of\ residuals + \lambda}$$

$$Gain = Similarity\ score(leafes) - Similarity\ score(root)$$

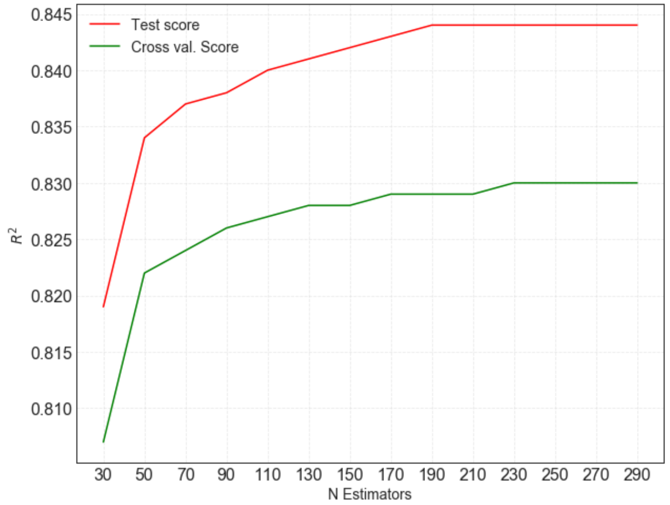
4.3 XGBoost

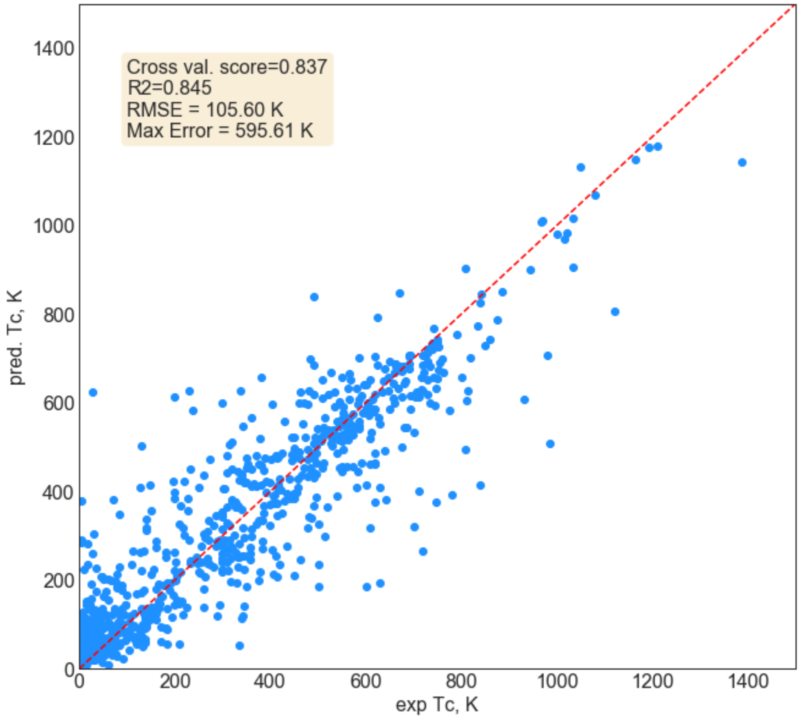
Hyperparameters tuning and model perfomance

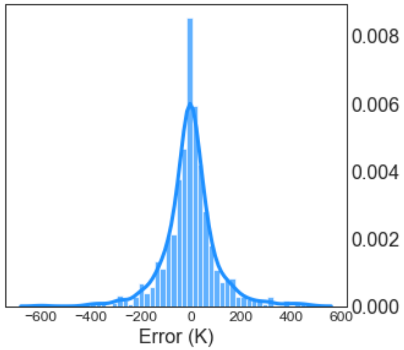
Tree depth optimization

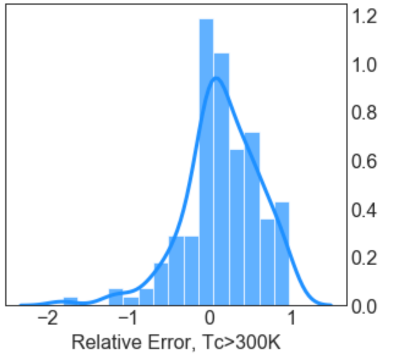


Number of Trees optimization









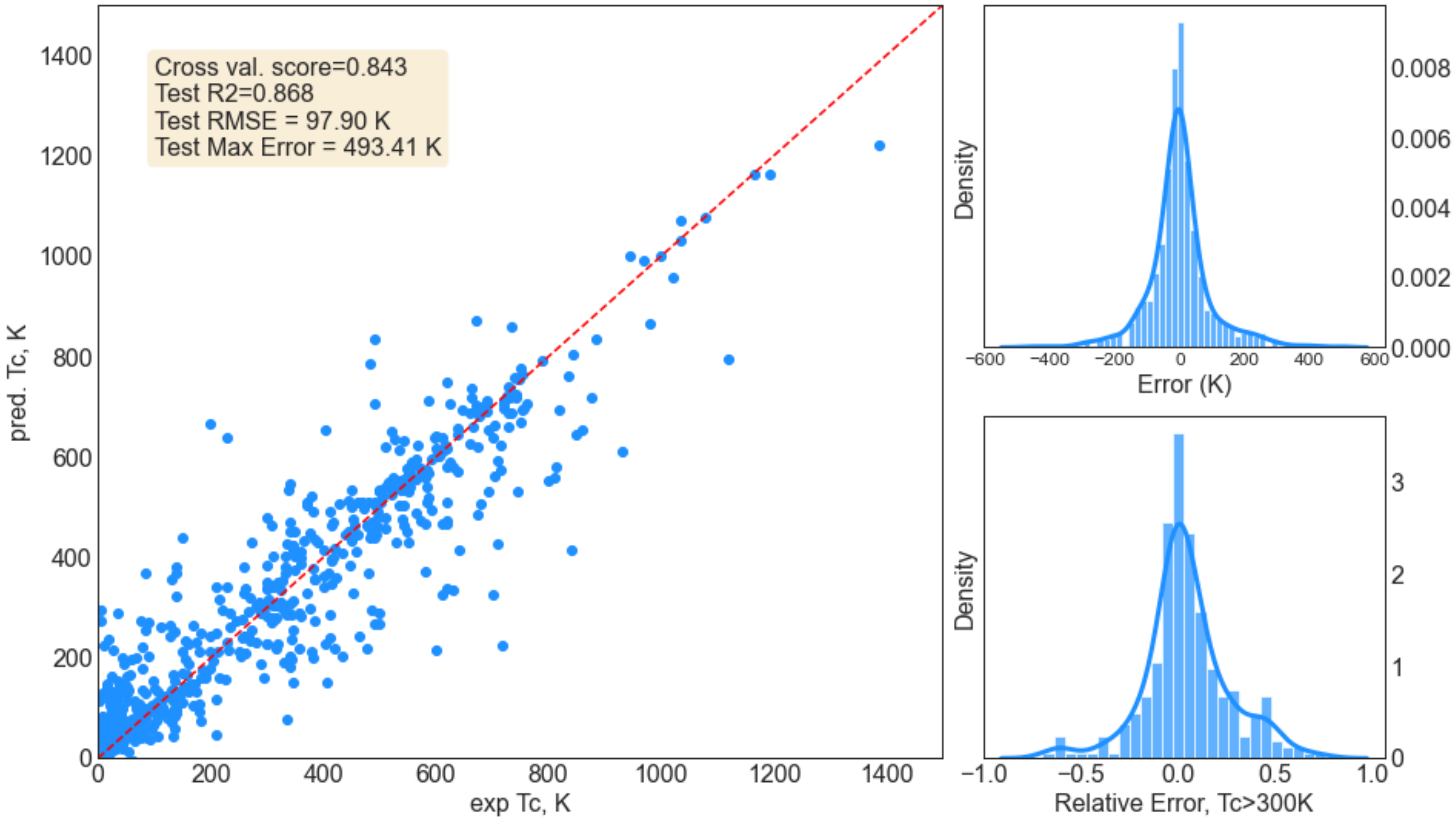
Metrics	Score
R^2 (test)	0.839 ↑
R^2 (cross.val)	0.878 ↑
RMSE (K)	93.85 ↑
Max Error (K)	463.34 ↓

4.4 LightGBM

Hyperparameters tuning and model perfomance

```
param_grid_lgb = {
    'learning_rate' : [0.01, 0.1],
    'n_estimators' : [500, 1000, 1500],
    'max_depth' : [10, 20],
    'num_leaves' : [100]
}
```

Metrics	Score
R^2 (test)	0.843 ↑
R^2 (cross.val)	0.868 ↑
RMSE (K)	97.90 ↑
Max Error (K)	493.41 ↓

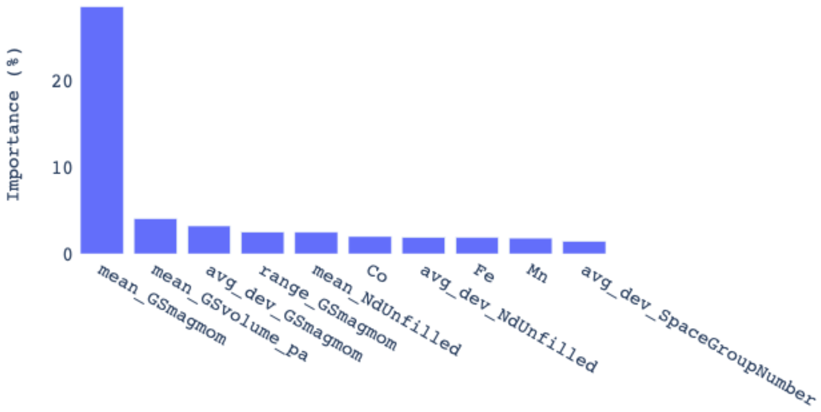


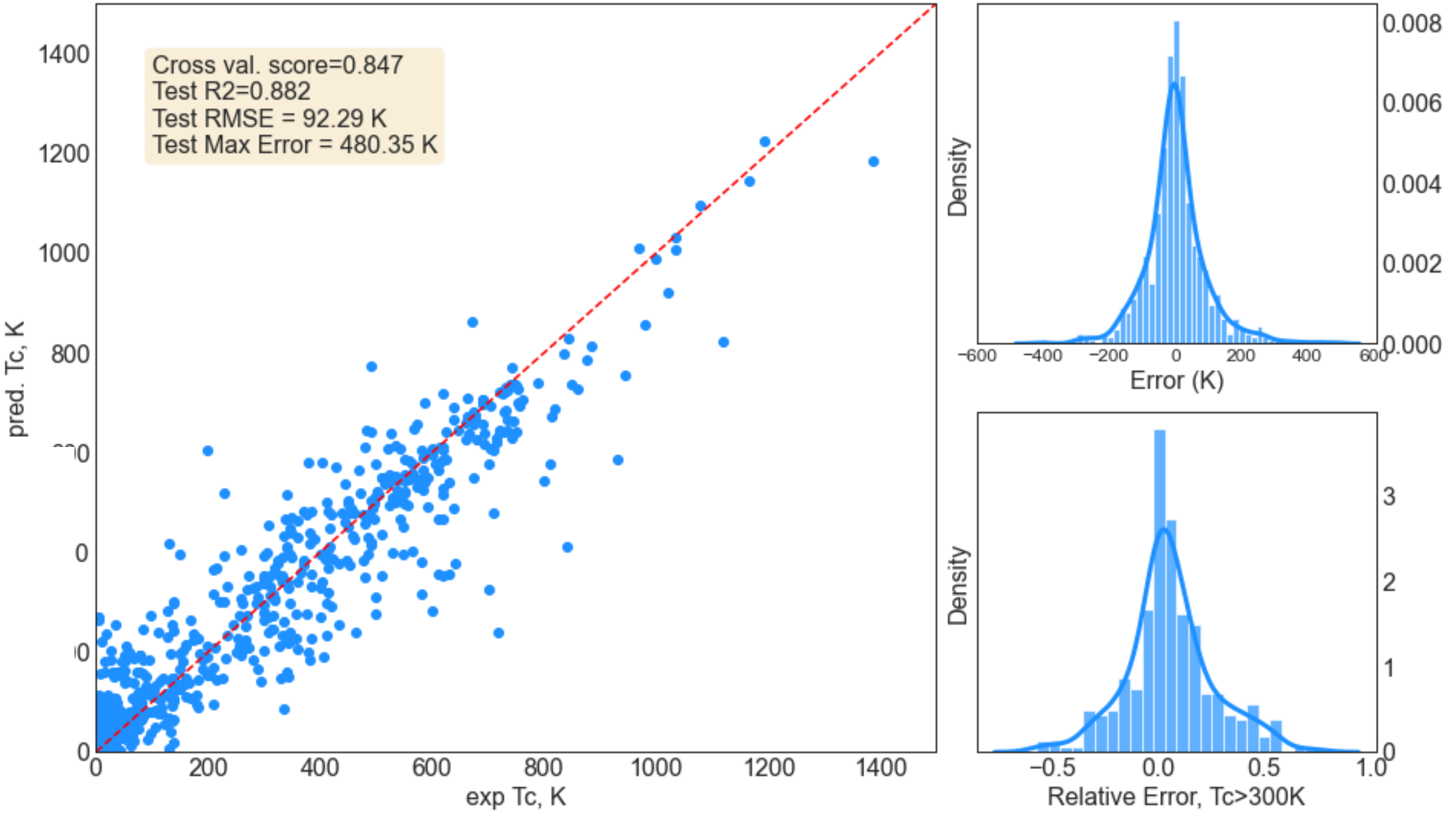
4.5 CatBoost

Model perfomance

Metrics	Score
R^2 (test)	0.847 ↑
R^2 (cross.val)	0.882 ↑
RMSE (K)	92.29 ↑
Max Error (K)	480.35 ↓

Feature by importances

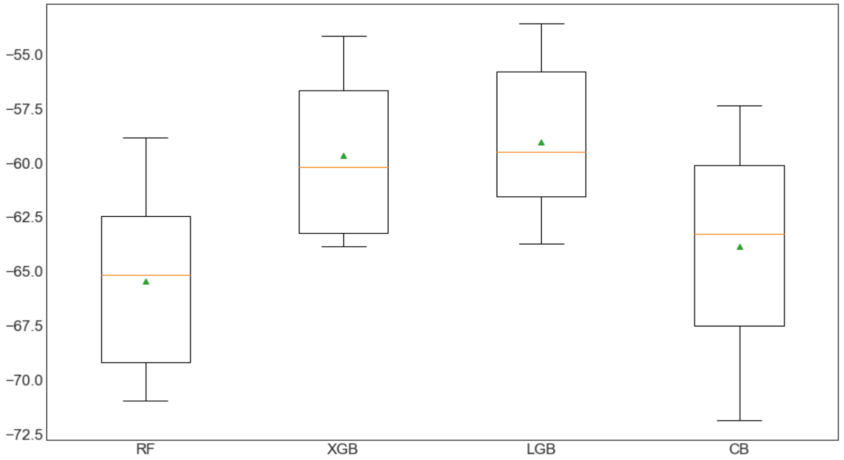


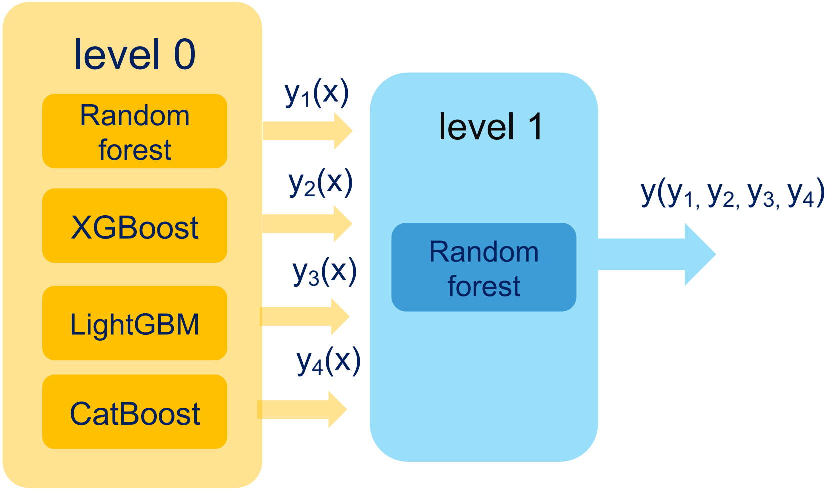


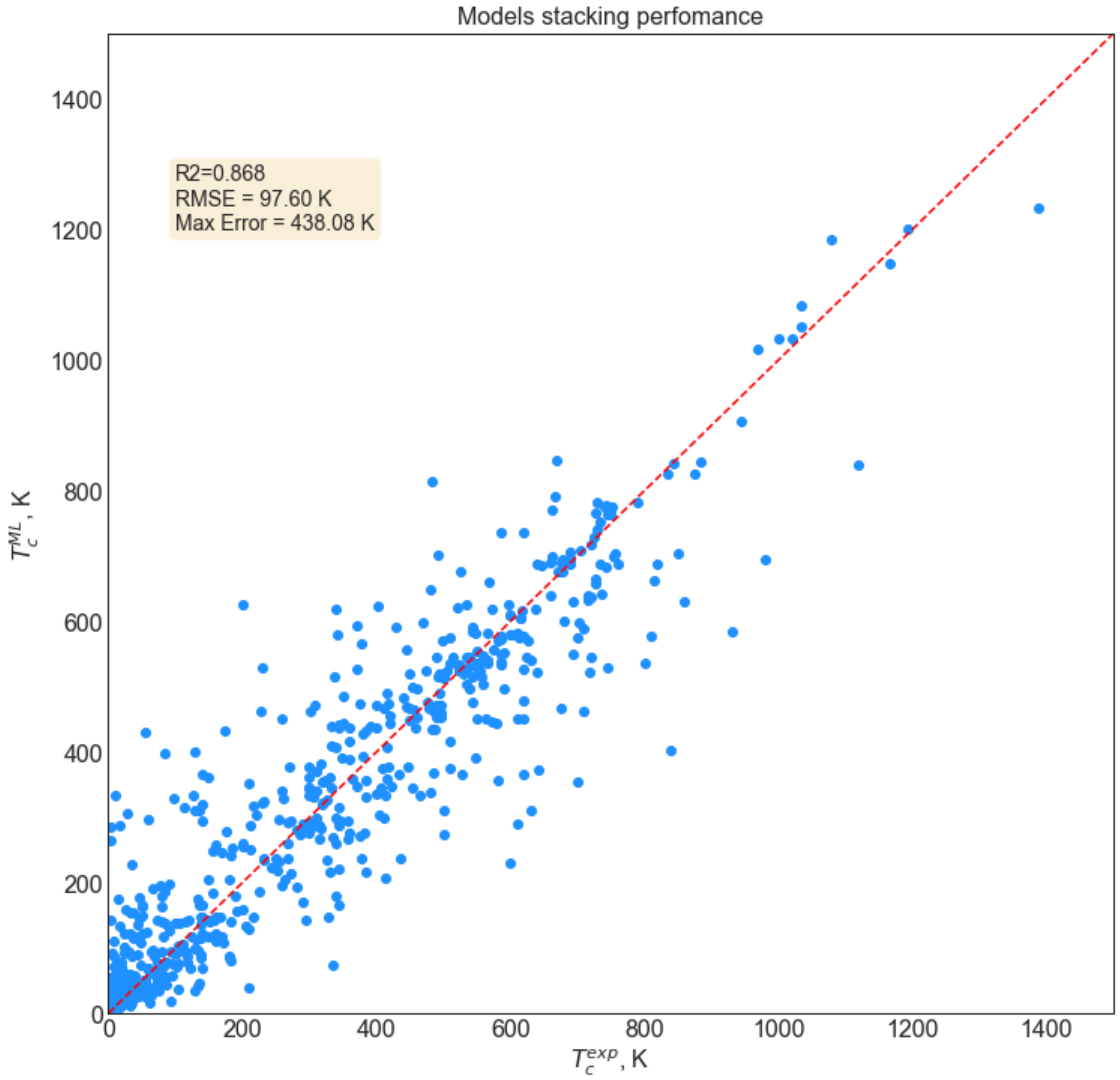
4.6 Models Stacking

Model perfomance

A box-and-whisker plot is then created comparing the distribution negative MAE scores for each model.



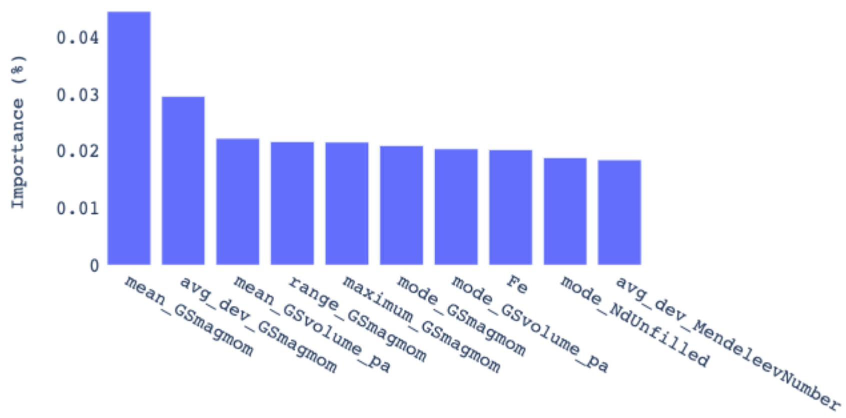




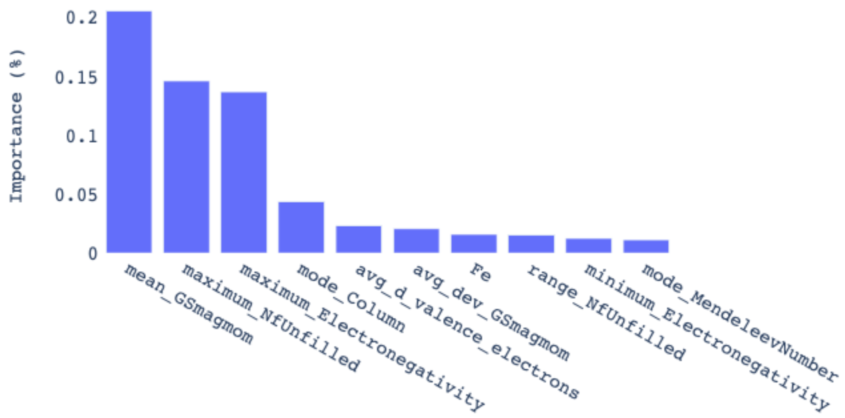
5. Feature Importance

Physical insights from trained models

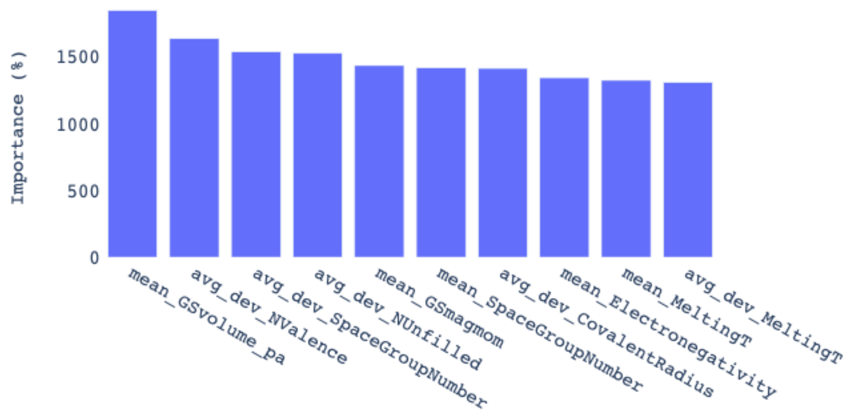
RandomForest



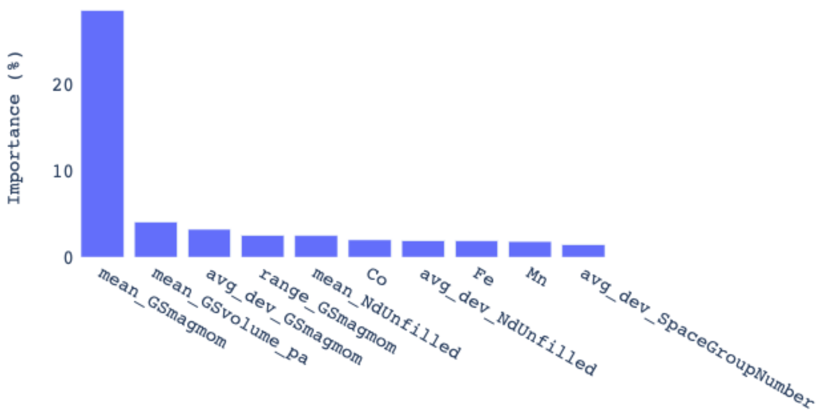
XGBoost



LightGBM



CatBoost



5. Summary of the best models

All the results on a single slide

Metrics	KRR	RF	XGB	LGB	CB	RF(ref.)	KRR(ref)
R^2 (test)	0.796	0.837	0.839	0.843	0.847	0.87	0.72
R^2 (cross.val)	0.818	0.859	0.878	0.868	0.882	0.81	-
RMSE (K)	114.73	101.04	93.85	97.90	92.29	57	-
Max Error (K)	637.59	468.08	463.34	493.41	480.35	-	-

5. Futher steps

Possibilites for this project

What can be done next:

1. Split the problem into the classification and regression (medium)
2. Increase the size of the dataset (hard)
3. Include structural information (serious drawback) (hard)

Meet Our wonderful team

Faridun Jalolov

Olajide Olubovale

Opeyemi Isaac

Dmitry Volkov