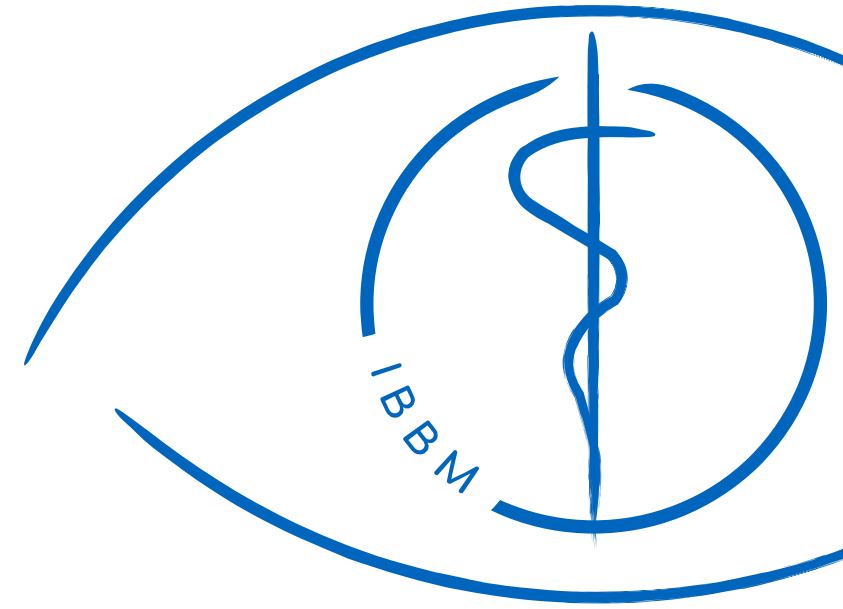
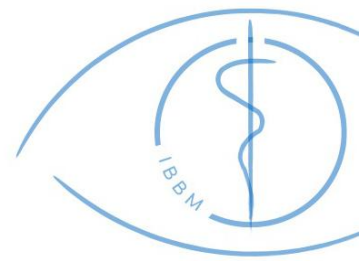


Update  
06.10.2020

---

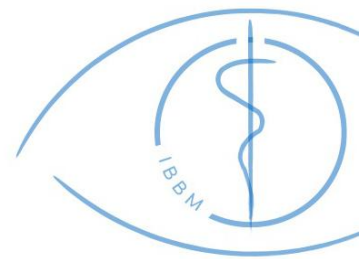
Ing. John LaMaster





# Results of RF

- Data:
  - Short TE
  - 10.000 spectra
  - 70/30 split
  - Cropped to [1000:1256]
- RandomForestRegressor
  - criterion: 'mse'
  - n\_estimators = 100 / 300
- Good R2 scores
- Okay accuracy - consistent problem with Cho accuracy in test set



# Discussion

- Choline accuracy - wildcard
  - As I re-split the train/test sets. the exploding choline acc changes between the sets
  - Only choline is affected - why?
- Suggestions?

## • Random Forest - Error Table

Trees	Phase	Cho	Cre	Naa	Total	R2
300	Train	0.786209212	1.00	0.891662835	0.8926	0.990872371
	Test	-353.5050962	1.00	0.735635941	-117.2565	0.932247276
100	Train	0.800089063	1.00	0.896737959	0.8989	0.990409992
	Test	-366.8325316	1.00	0.73043075	-121.7007	0.931319104
10	Train	0.784829446	1.00	0.885101507	0.8900	0.984146803
	Test	-336.2962249	1.00	0.748288326	-111.5160	0.918872949
5	Train	0.805176881	1.00	0.870783054	0.8920	0.977605852
	Test	-440.8599049	1.00	0.715611491	-146.3814	0.899829768

- \*\*Note: 1 - % Error (referred to above as accuracy)