

Importance Of Feature Selection in ML Models or: How I Learned to Start Worrying and Love Ensembles



Sankalp Gilda
Graduate Student, Dept. of Astronomy
University of Florida

Feature Selection...

- 1) **What** is it?
- 2) **Why** do we need it?
- 3) **How** do we do it?
- 4) Does it **help**?

Feature Selection...

- 1) **What** is it?
- 2) **Why** do we need it?
- 3) **How** do we do it?
- 4) Does it **help**?

Feature Selection...

1) ~~What~~ is it?

→ 2) **Why** do we need it?

3) **How** do we do it?

4) Does it **help**?

Because this.

THIS IS YOUR MACHINE LEARNING SYSTEM?

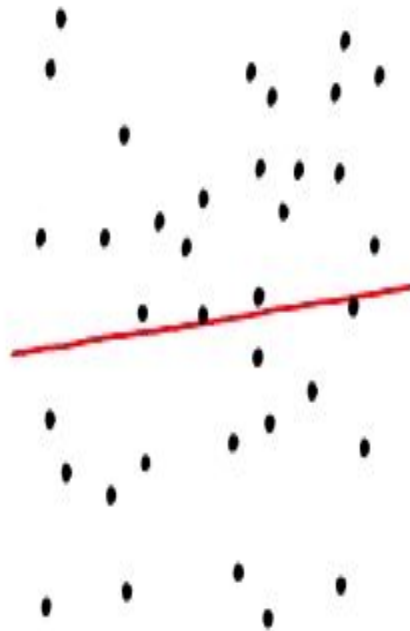
YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.

WHAT IF THE ANSWERS ARE WRONG?

JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.



And this.



$$R^2 = 0.06$$



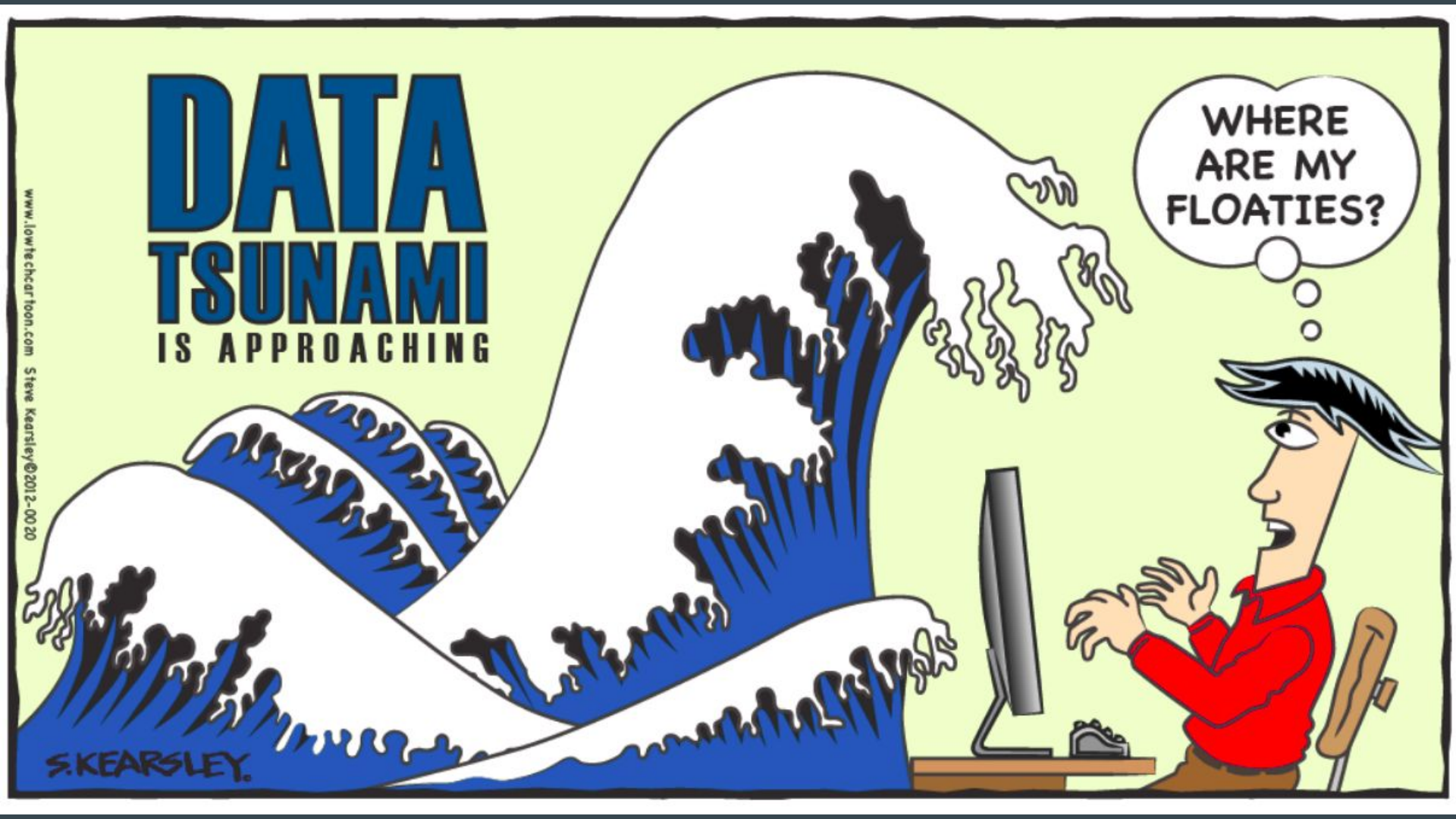
REXTHOR, THE DOG-BEARER

I DON'T TRUST LINEAR REGRESSIONS WHEN IT'S HARDER TO GUESS THE DIRECTION OF THE CORRELATION FROM THE SCATTER PLOT THAN TO FIND NEW CONSTELLATIONS ON IT.

DATA TSUNAMI IS APPROACHING

WHERE
ARE MY
FLOATIES?

S. KEARSLEY



Feature Selection...

1) ~~What~~ is it?

2) ~~Why~~ do we need it?

→ 3) **How** do we do it?

4) Does it **help**?

Feature Selection...

1) ~~What~~ is it?

2) ~~Why~~ do we need it?

→ 3) ~~How~~ do we do it? How did I do it?

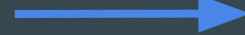
4) Does it **help**?

Recursive
Backward
Elimination

Random Sub-Space
Selection

Backward
Elimination

Predictions from
Stack



Bagging



Hyperparameter
optimization (CV)



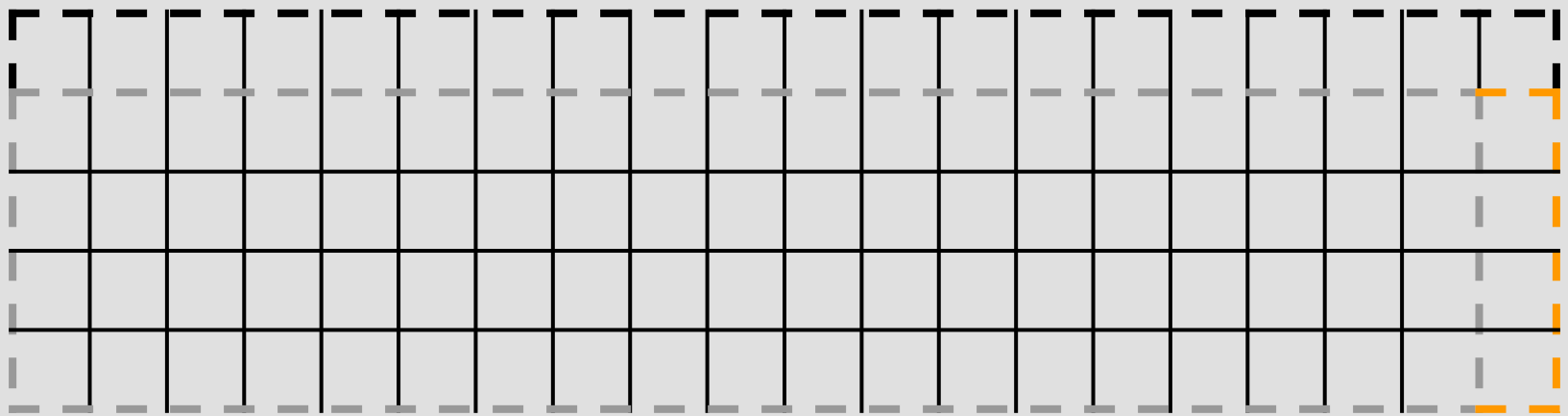
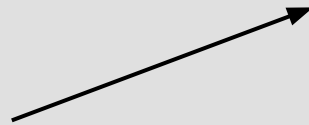
Predictions from
KNNs



Aggregation of
Features



Header

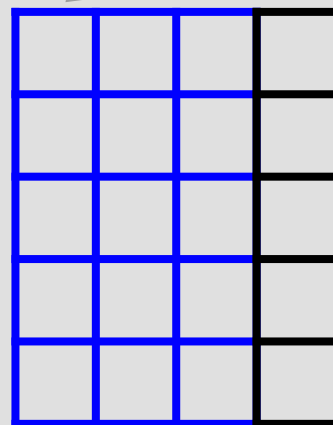
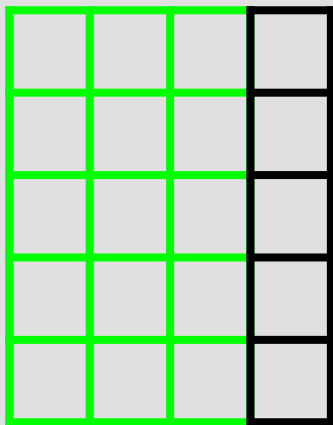
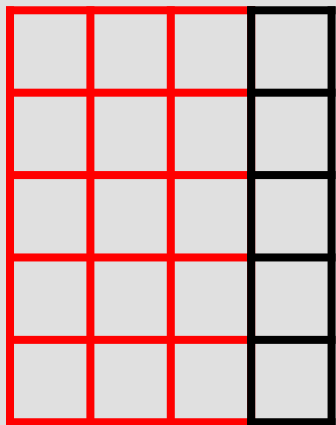
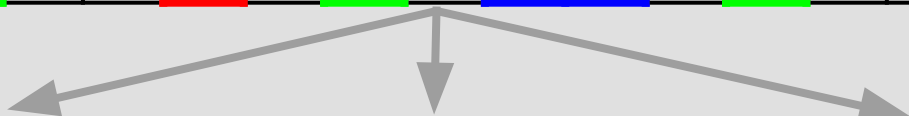
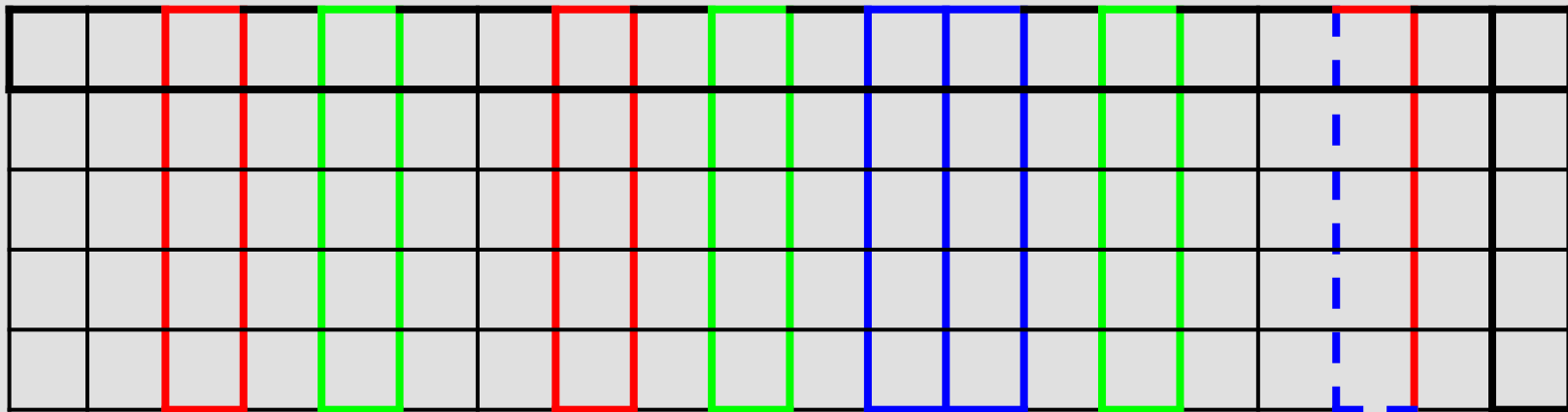


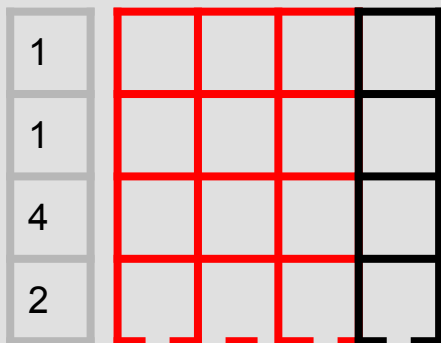
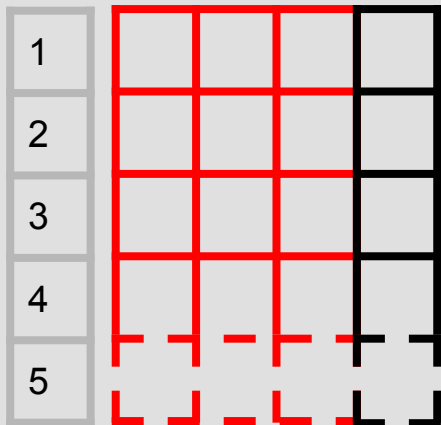
Predictors/ Features



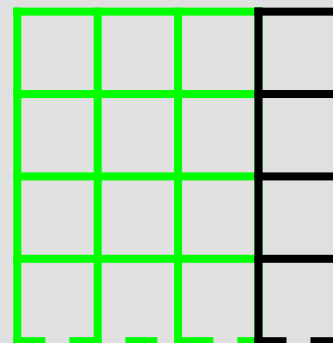
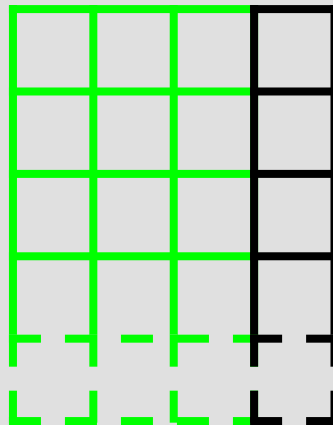
Target



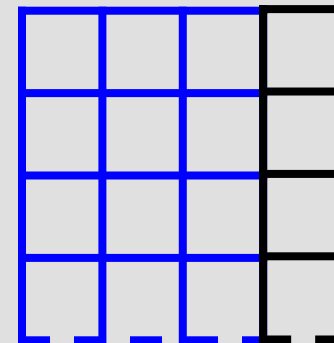
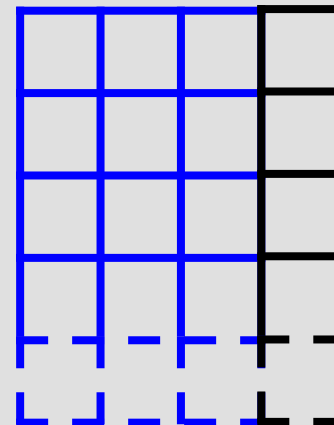




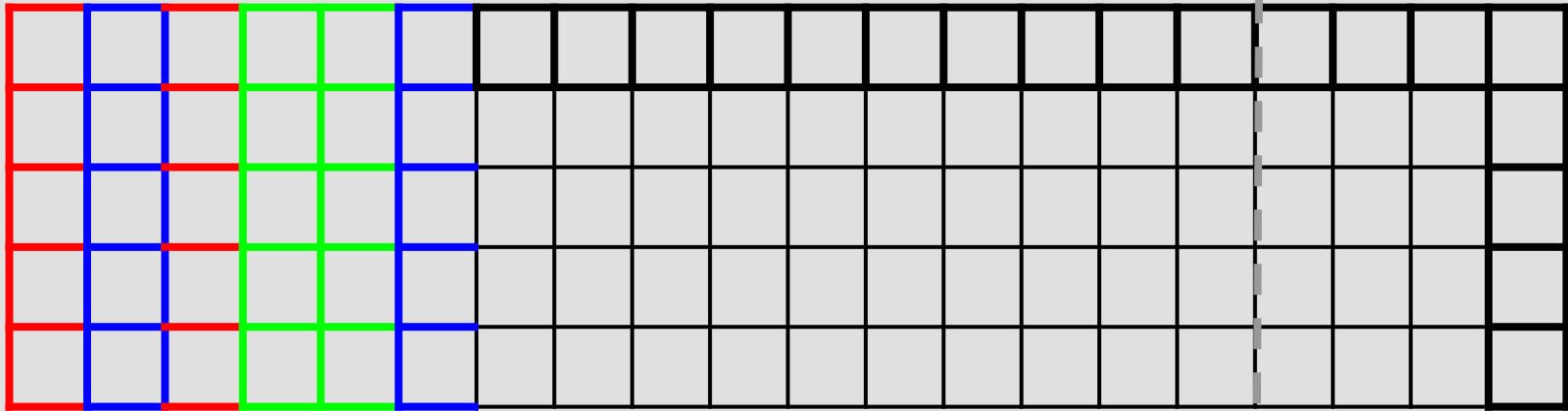
Predictions_0



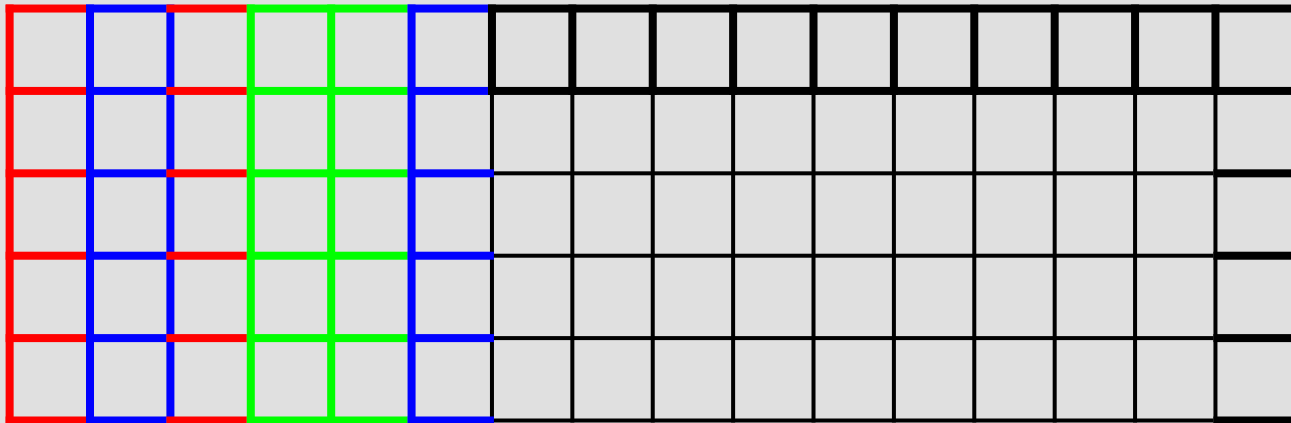
Predictions_1



Predictions_2



→ <Insert Stack Here>



Feature Selection...

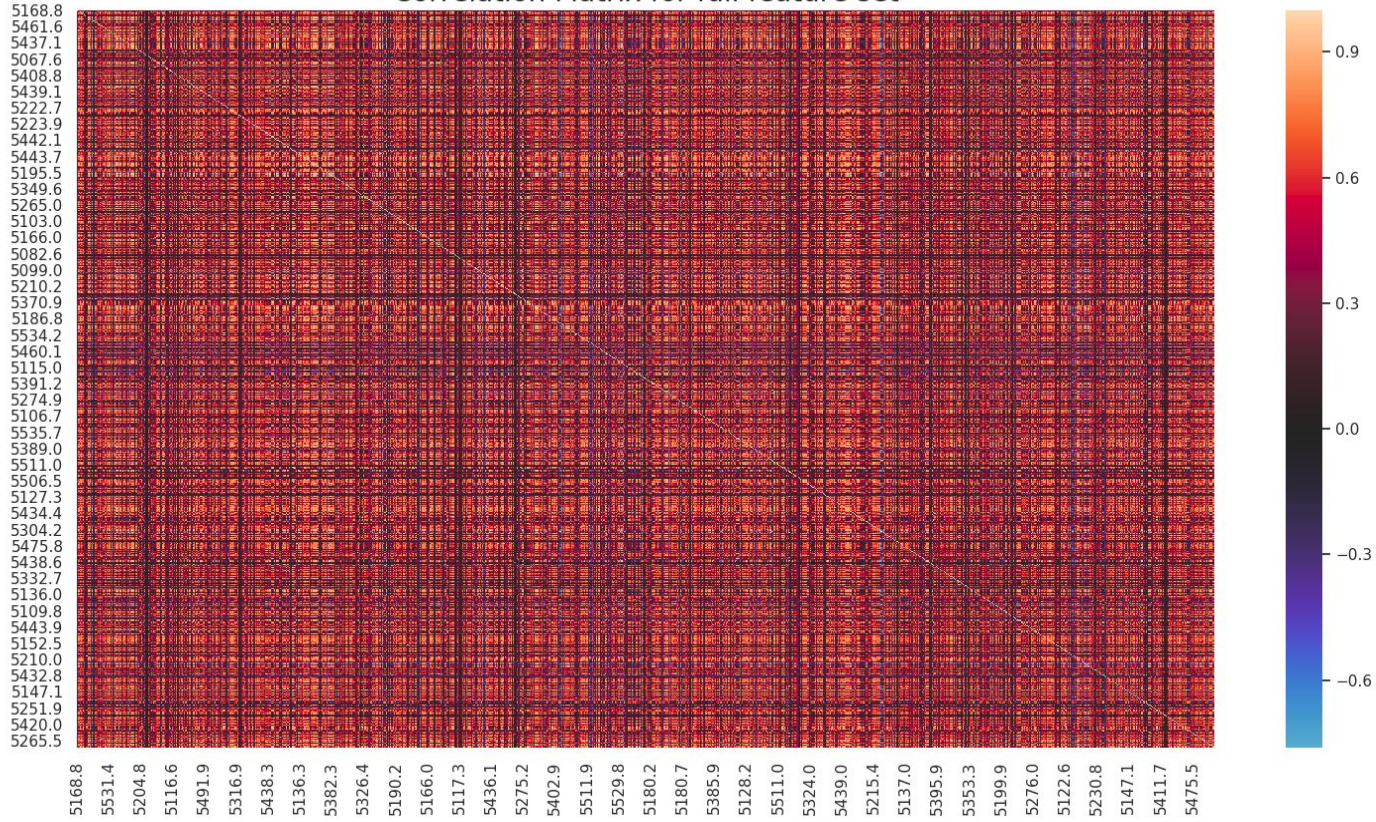
1) ~~What~~ is it?

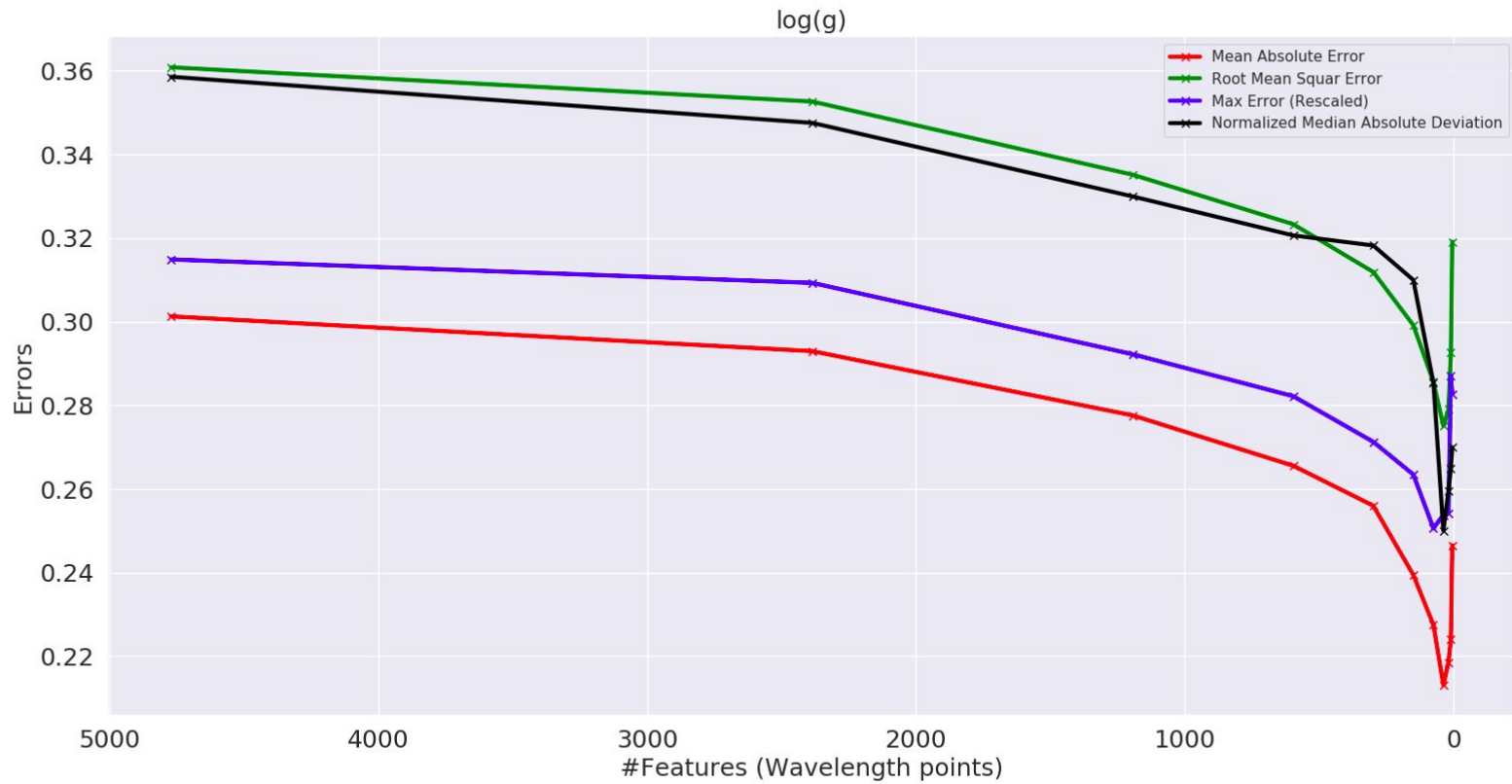
2) ~~Why~~ do we need it?

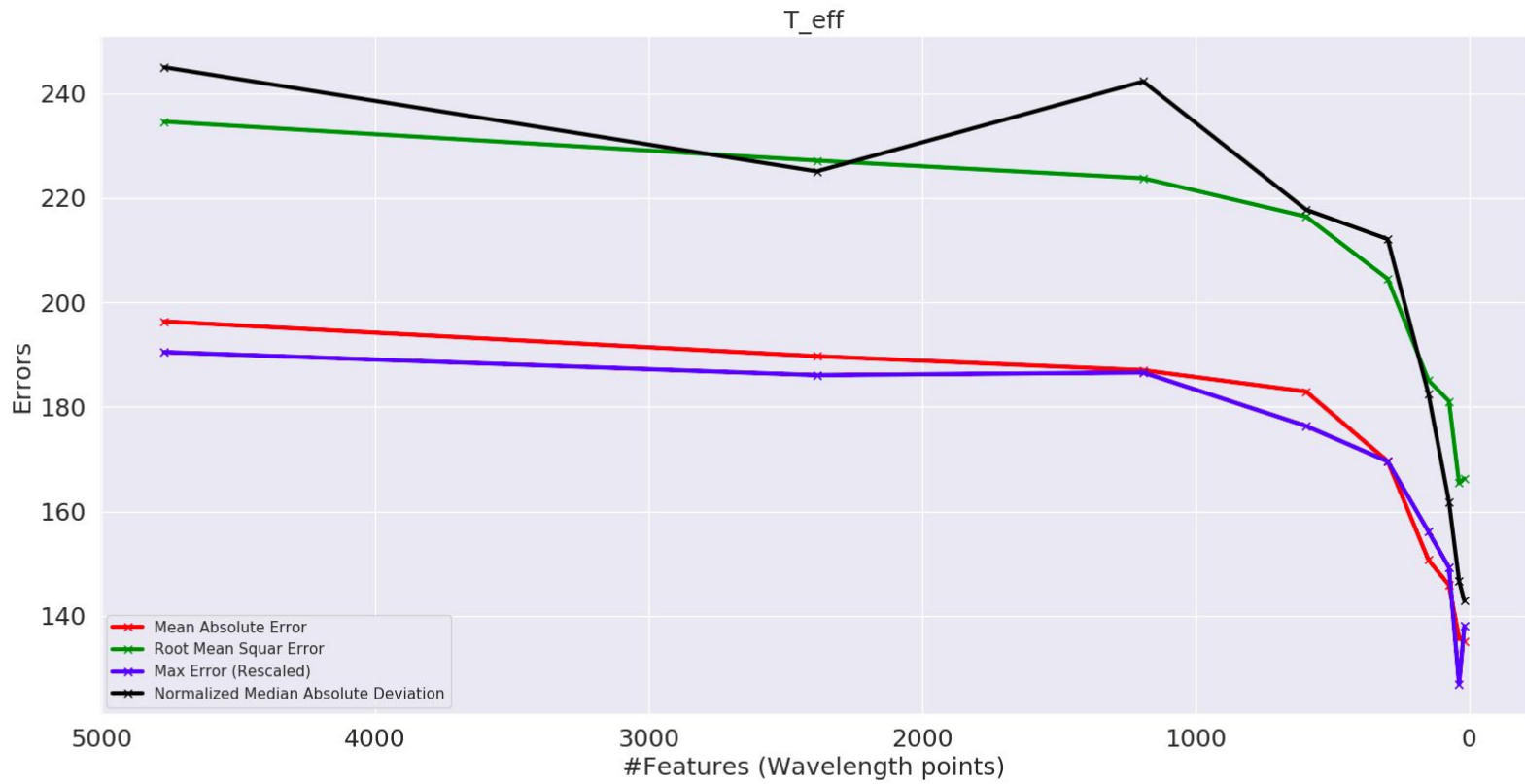
3) ~~How~~ do we do it? ~~How did I do it?~~

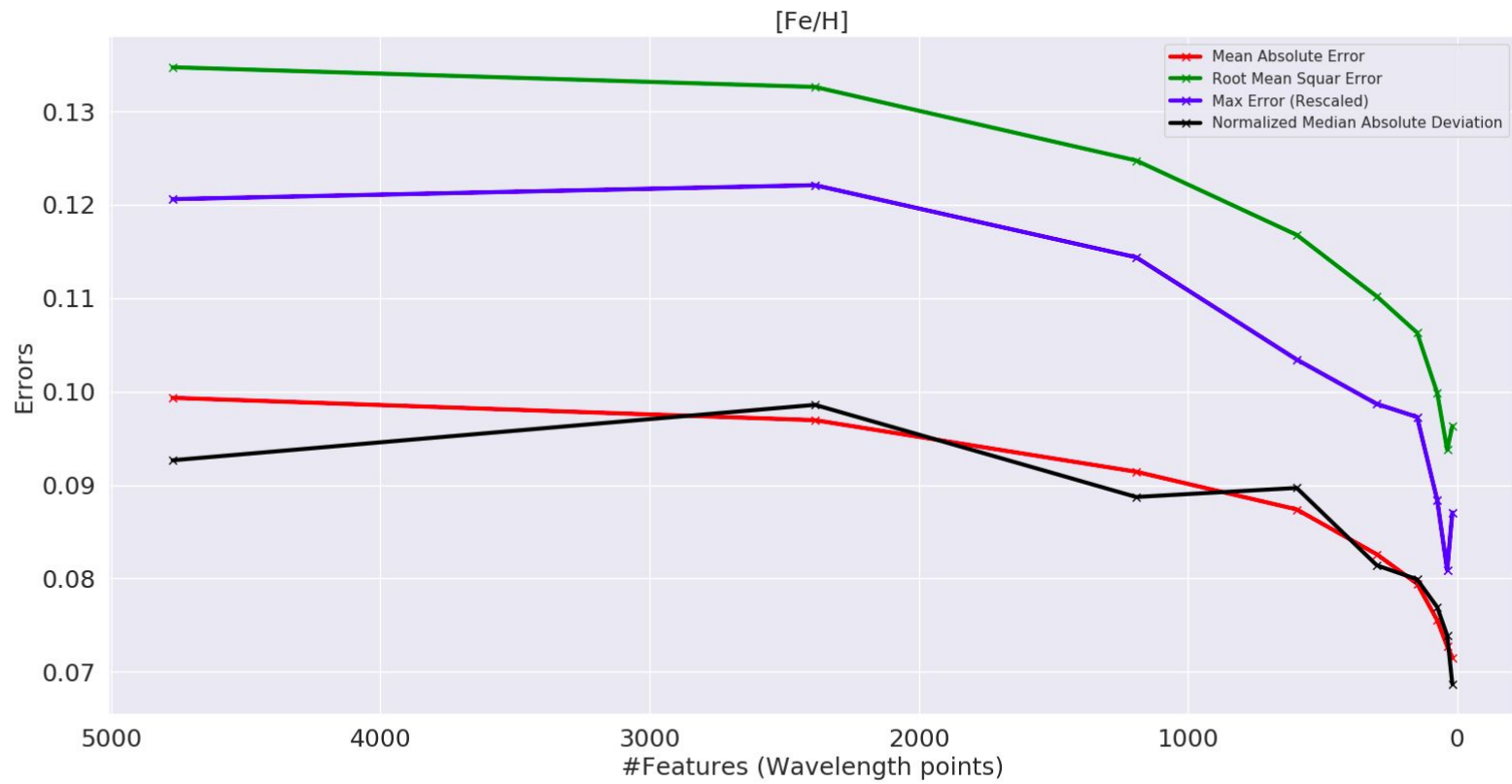
→ 4) Does it **help**?

Correlation Matrix for full feature set









Feature Selection...

~~1) **What** is it?~~

~~2) **Why** do we need it?~~

~~3) **How** do we do it? How did I do it?~~

~~4) Does it **help**?~~ ✓

Feature Selection...

1) ~~What~~ is it?

2) ~~Why~~ do we need it?

3) ~~How~~ do we do it? ~~How did I do it?~~

4) ~~Does it help?~~ ✓

→ 5) Is it **robust**?

Robustness



Feature Selection...

1) ~~What~~ is it?

2) ~~Why~~ do we need it?

3) ~~How~~ do we do it? ~~How did I do it?~~

4) ~~Does it help?~~ ✓

→ 5) ~~Is it robust?~~ ✓

What's Next?

- **Move to a Probabilistic Framework (80%)**
- **Clean up code, and free up hard-coded params (80%)**
- **Create User Guide (50%)**
- **Assess performance against resolution, SNR, #samples**
- **Post to GitHub by end of Nov.**

Questions?

Please feel free to get in touch if you have any questions, concerns or suggestions!

email: s.gilda@ufl.edu