



September 20th-22nd, 2022
MTEch Users Conference





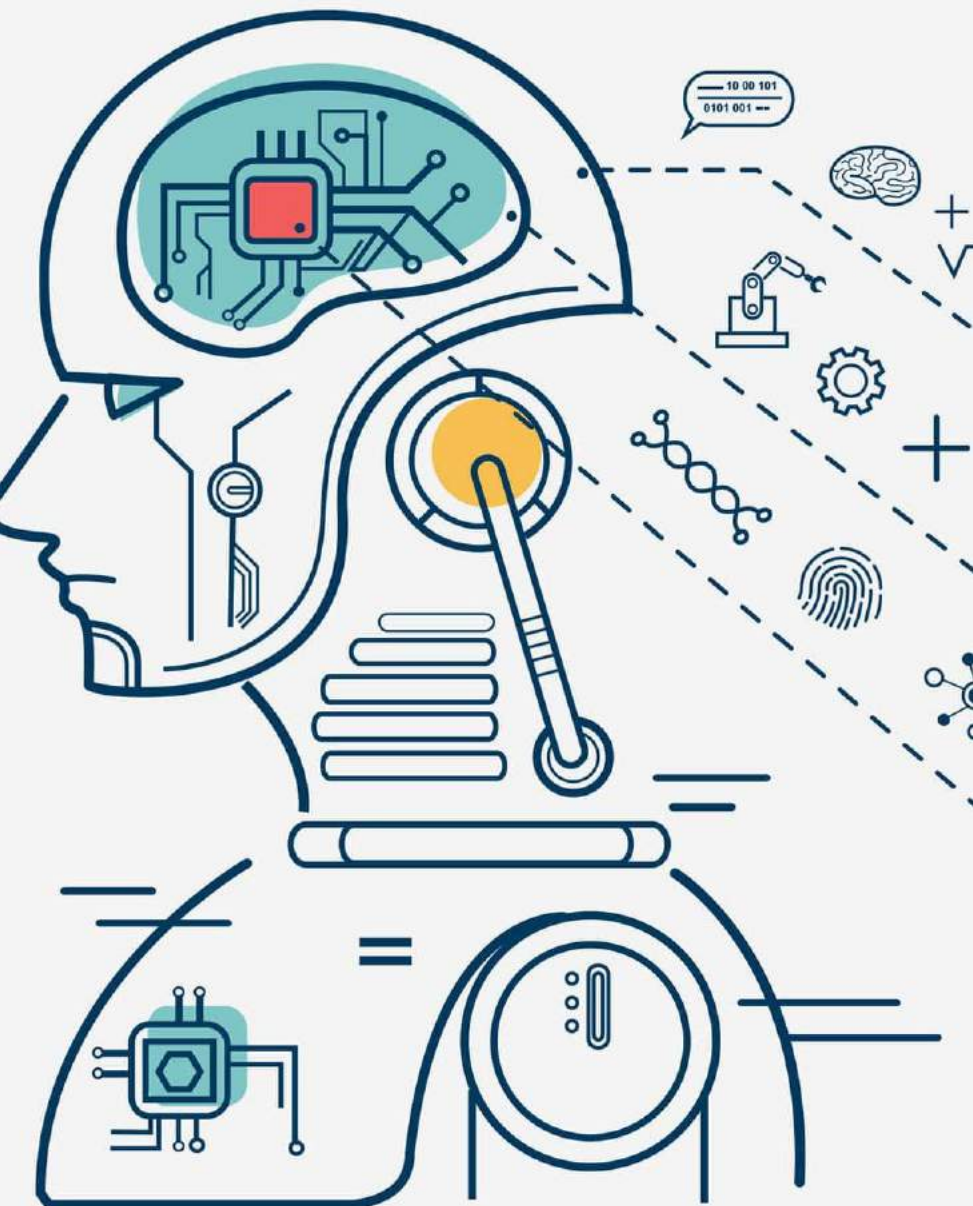
September 20th-22nd, 2022
MTEch Users Conference



Evan Sadlon – Data Science Manager

Data Science Detective Work





Goals for Talk

Feature Engineering

- Domain Expertise
- Automation tools

Feature Importance

- Calculation
- Selecting

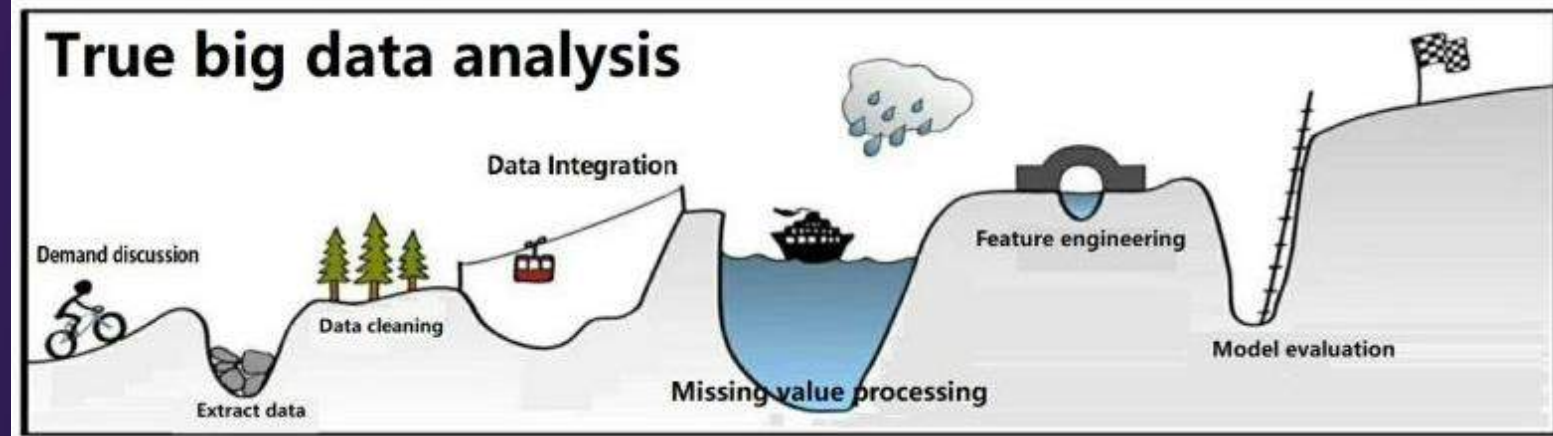
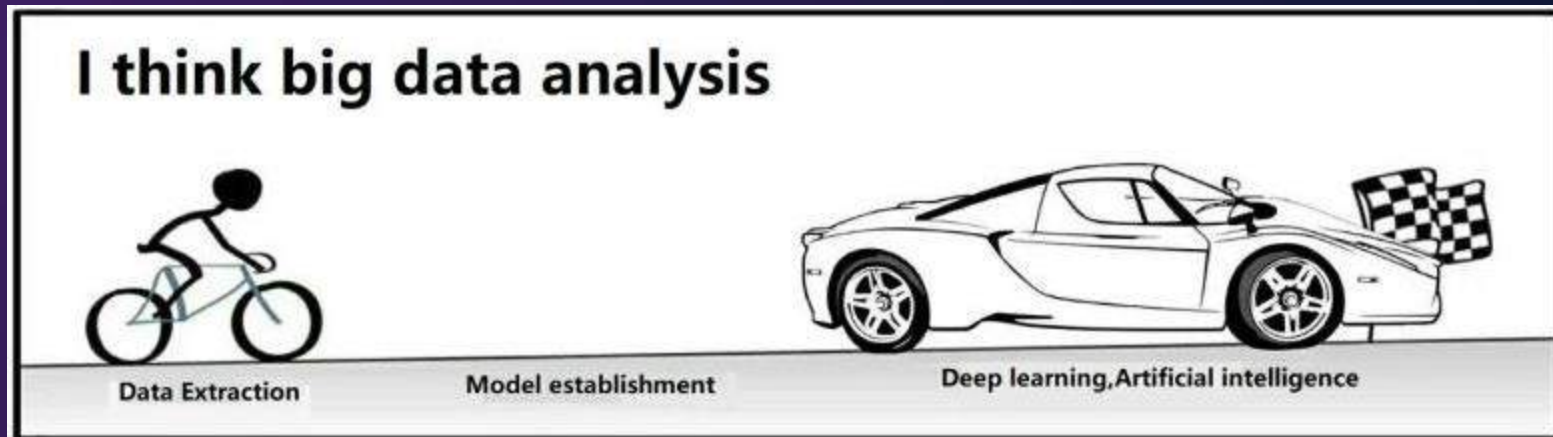
Modeling

- Interpreting
- Taking Action
- Knowing when to ignore the machine



September 20th-22nd, 2022
MTech Users Conference

What is Machine Learning



September 20th-22nd, 2022
MTech Users Conference

Data cleaning

Invalid data

Zero or negative values in...

- - *Weekly Field Transaction Feed Consumed*
- - *Parent Age*
- - *Initial Weight*

Sum of *Cull*, *Mortality*, and *Head Sold* not equaling *Head Placed*

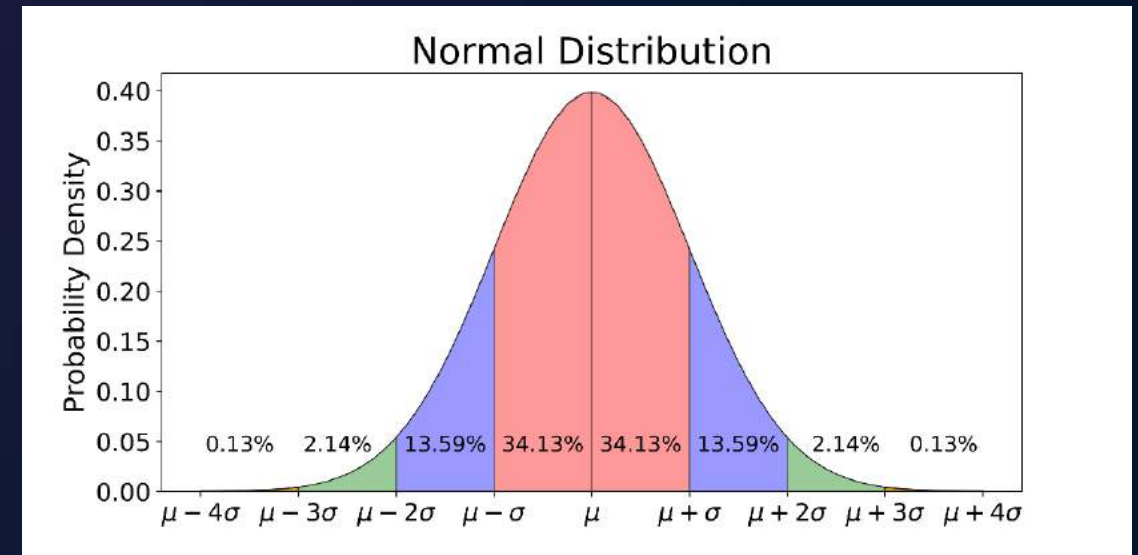
Some weekly *Field Transaction Weights* not increasing

Outliers

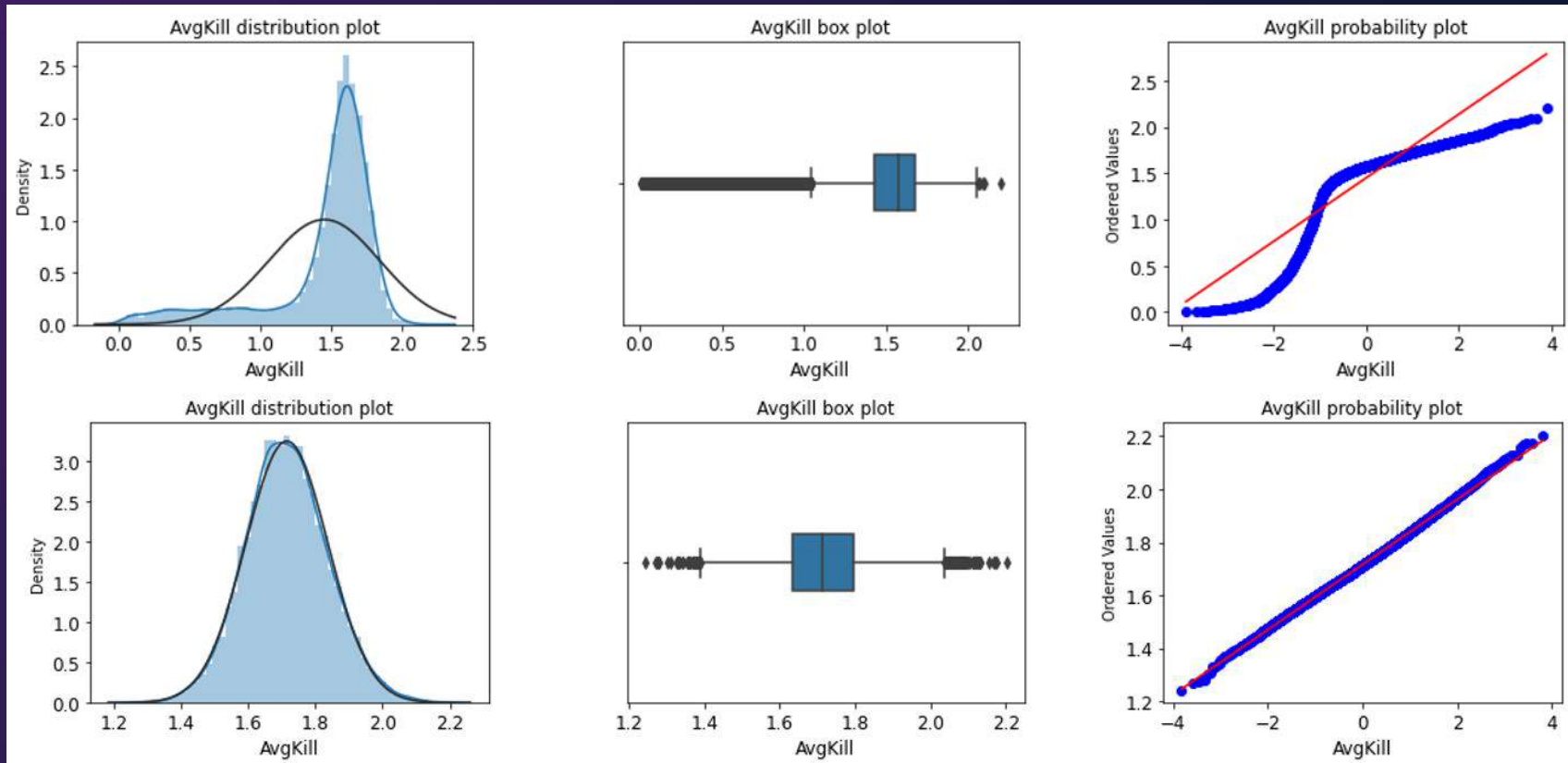
- *FCR* value over 3.0
- *Initial Weight* under 30g and over 60g
- *Parent Age*
- *Weekly Field Transaction Weights*

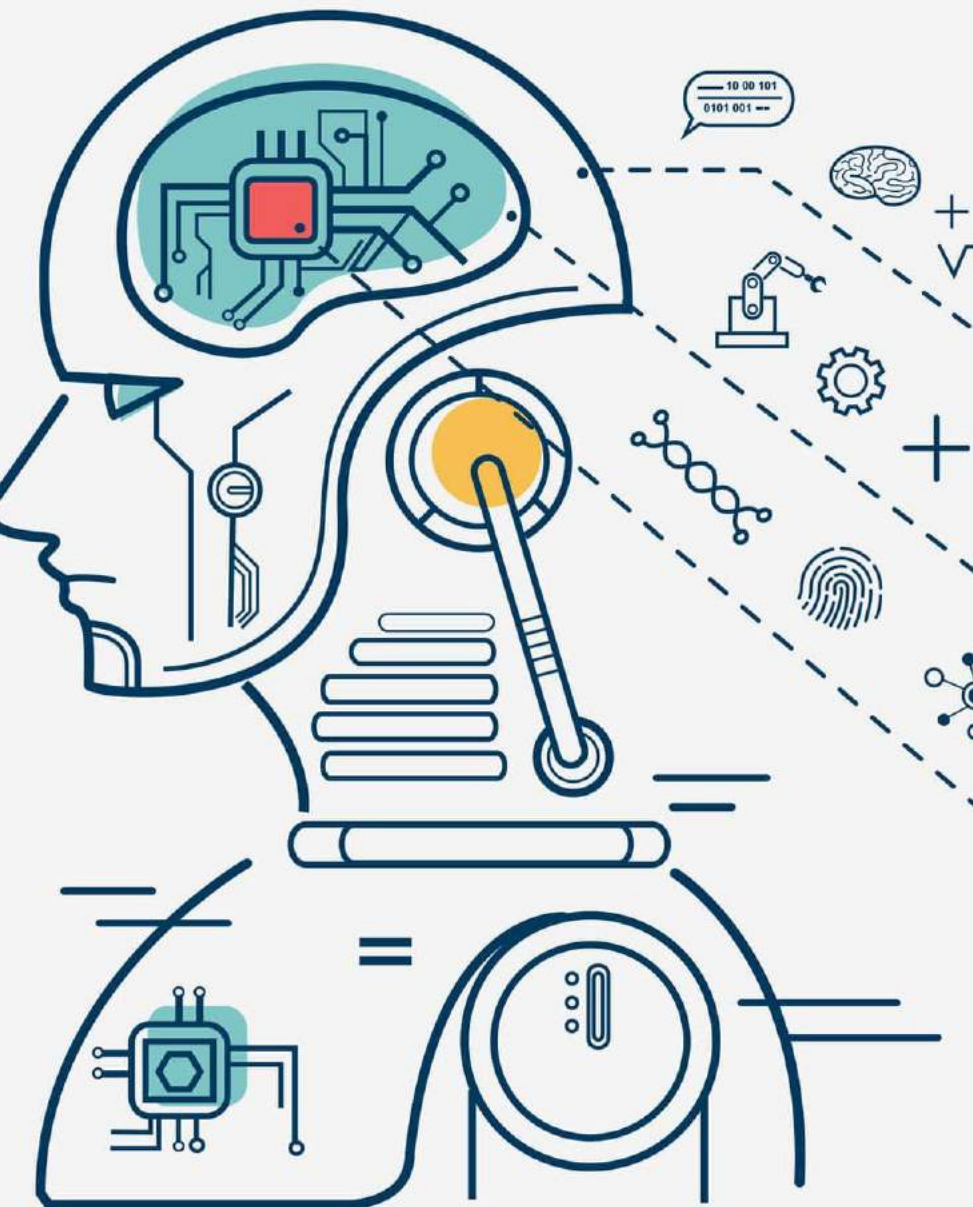
Missing values columns

- Protein Tech Advisor IRN
- Protein Facility Plants IRN
- FDSA Formula Number Assignment IRN



What is Machine Learning





Goals for Talk

Feature Engineering

- Domain Expertise
- Automation tools

Feature Importance

- Calculation
- Selecting

Modeling

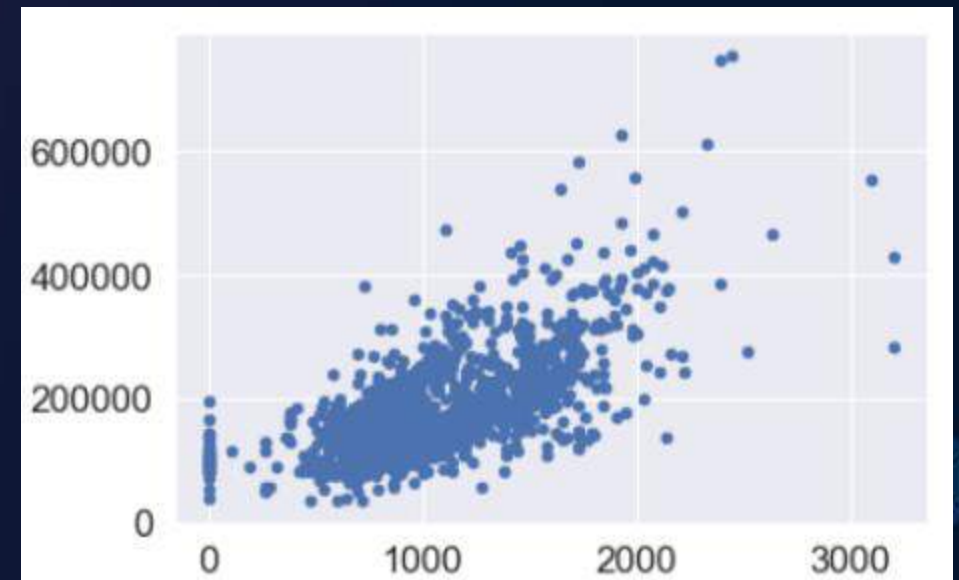
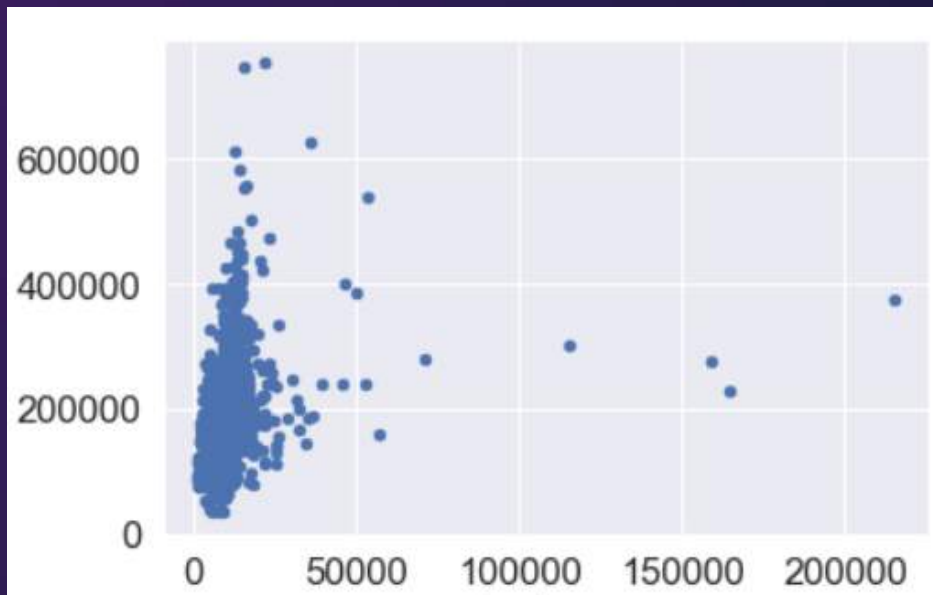
- Interpreting
- Taking Action
- Knowing when to ignore the machine



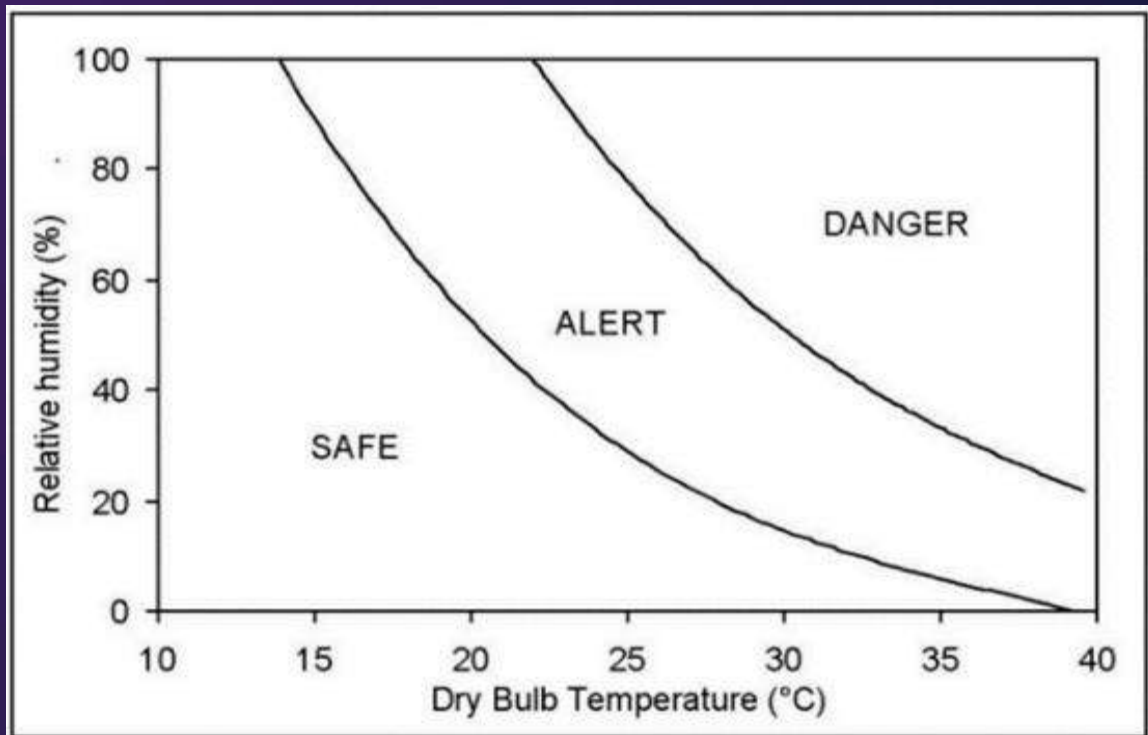
September 20th-22nd, 2022
MTech Users Conference

Feature Engineering

- The act of using domain knowledge by a subject matter expert to make new datapoints that help the machine learn trends and patterns



Feature Engineering – Domain Expertise



- Stand on the shoulders of giants
- Talk to as many different people as possible
- But don't be afraid to ignore someone's lifehack



Feature Engineering – Domain Expertise

- **AET – By Week and Life of Flock Trend**
 - Stability
 - Trend
 - Across house
- **Curtain opening half to full house brood**
 - What age
 - What temperature when curtains opened
- **Chick receiving**
 - What temperature was house when chicks delivered
 - What outside temperature when chickens delivered



September 20th-22nd, 2022
MTech Users Conference

Feature Engineering - Automated

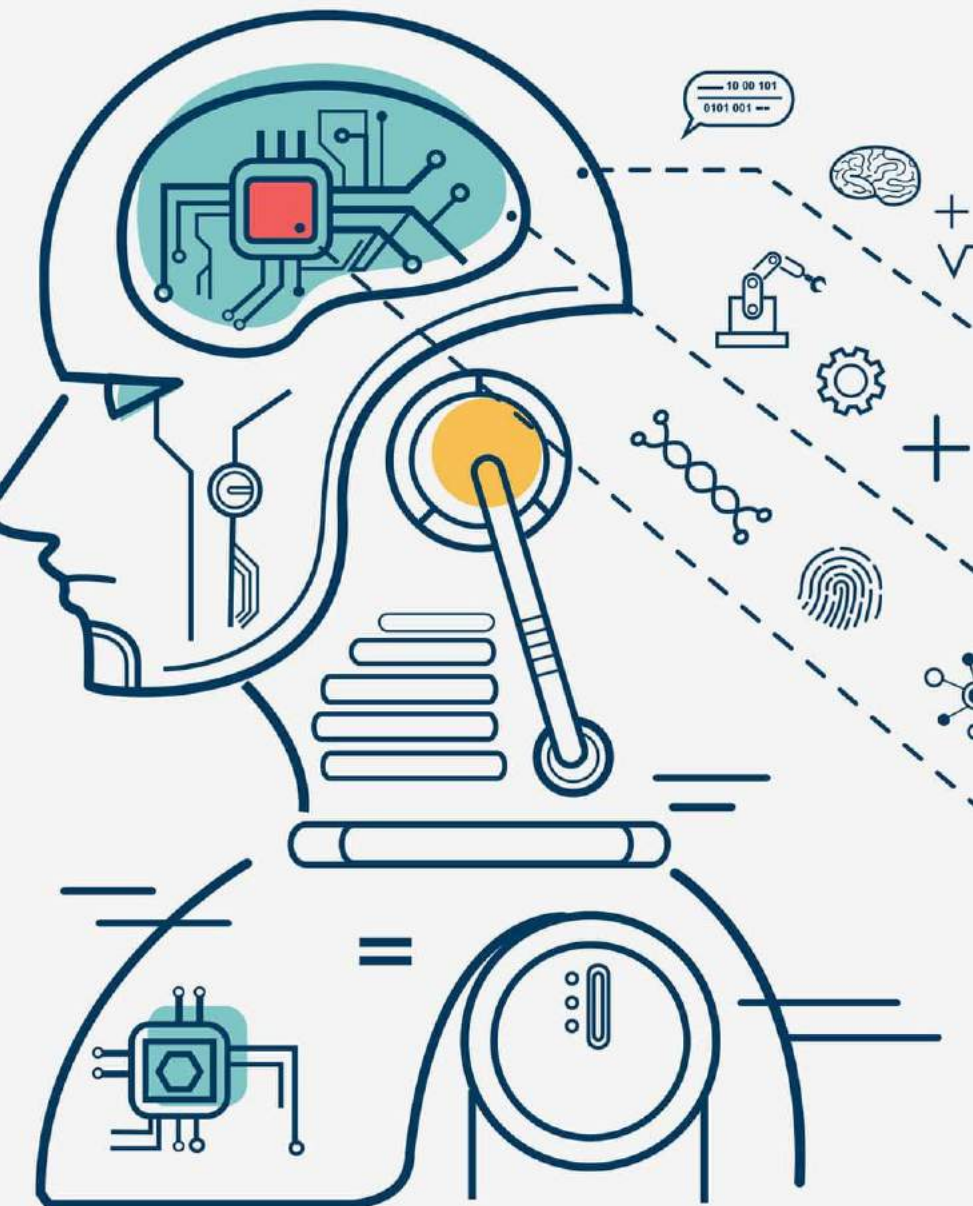
- This can be anything from combining multiple columns, to aggregating over past data. We can turn 6 columns into 600+ columns, but why and how do we choose?

SonarAge	SonarBirdWeight	Temperature1	Humidity1
0	0.06723	32.37010309	61.17938144
1	0.06394	32.14123711	57.95154639
2	0.07703	31.39896907	66.00618557
3	0.09541	30.95257732	59.99484536
4	0.11583	30.4	61.10824742
5	0.14294	29.92989691	66.39896907
6	0.19279	29.49896907	68.61958763
7	0.19248	29.32886598	71.83711134
8	0.25711	28.99072165	74.08659794
9	0.28236	28.61134021	72.83092784
10	0.33399	28.5556701	71.86185567
11	0.38925	28.12783505	70.71958763
12	0.43412	28.14742268	76.4



Age	Temp RMSE	Temp MaxLocation	Temp Linearness	Weight RMSE	Weight MaxLocation	Weight Linearness
0	32.3701	1	5.82E-14	26.2045267	0.020408163	1.27E-27
1	32.25587	0.5	0	26.3329871	0.020833333	5.07E-27
2	31.97279	0.333333333	0.1885935	26.4254113	0.021276596	5.52E-26
3	31.72081	0.25	0.01740217	26.529044	0.02173913	8.00E-25
4	31.46109	0.2	0.00097274	26.6365252	0.022222222	1.12E-23
5	31.21111	0.166666667	3.83E-05	26.7484486	0.022727273	1.42E-22
6	30.97231	0.142857143	1.47E-06	26.8528948	0.023255814	1.86E-21
7	30.77168	0.125	9.07E-07	26.9663198	0.023809524	2.14E-20
8	30.57892	0.111111111	2.02E-07	27.082379	0.024390244	2.17E-19
9	30.38789	0.1	2.30E-08	27.2024124	0.025	1.74E-18
10	30.22592	0.090909091	1.89E-08	27.3110784	0.025641026	1.91E-17
11	30.05667	0.083333333	2.66E-09	27.4223131	0.026315789	1.90E-16
12	29.91413	0.076923077	2.24E-09	27.5124857	0.027027027	1.77E-15





Goals for Talk

Feature Engineering

- Domain Expertise
- Automation tools

Feature Importance

- Calculation
- Selecting

Modeling

- Interpreting
- Taking Action
- Knowing when to ignore the machine



September 20th-22nd, 2022
MTech Users Conference



Feature Importance - Calculation

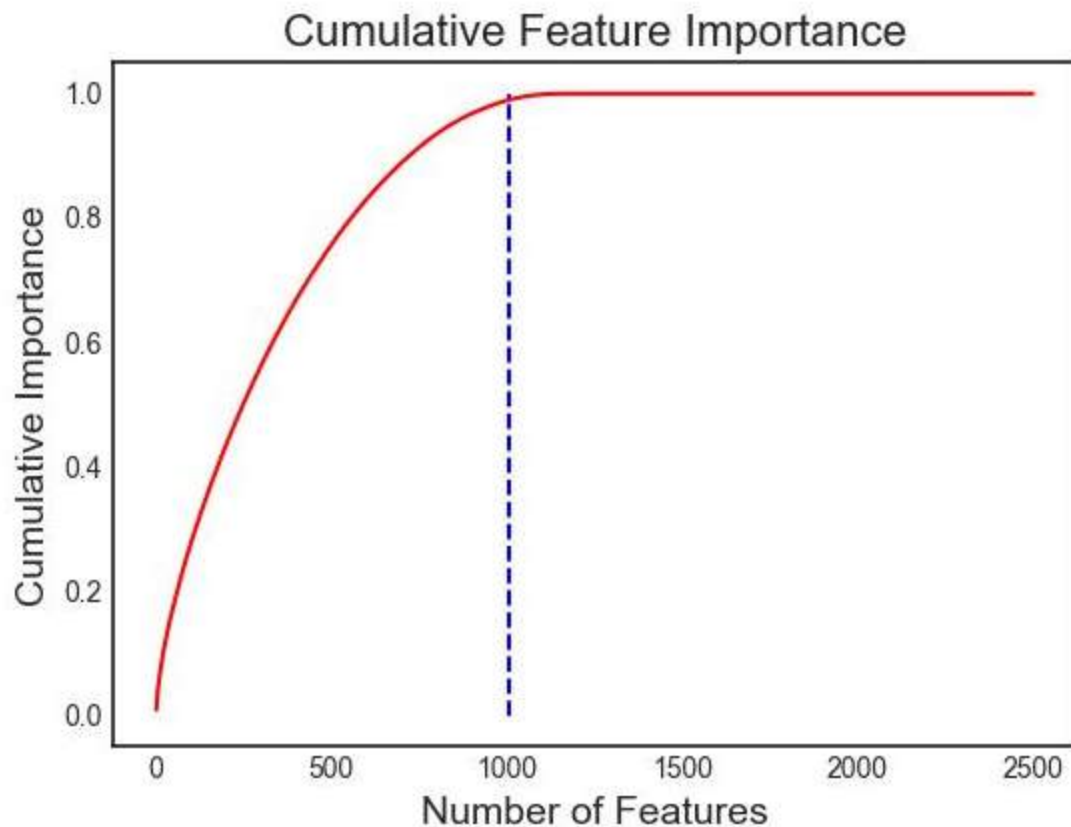


- “Interesting” data
- Accuracy gains/increases
- “Would you rather”



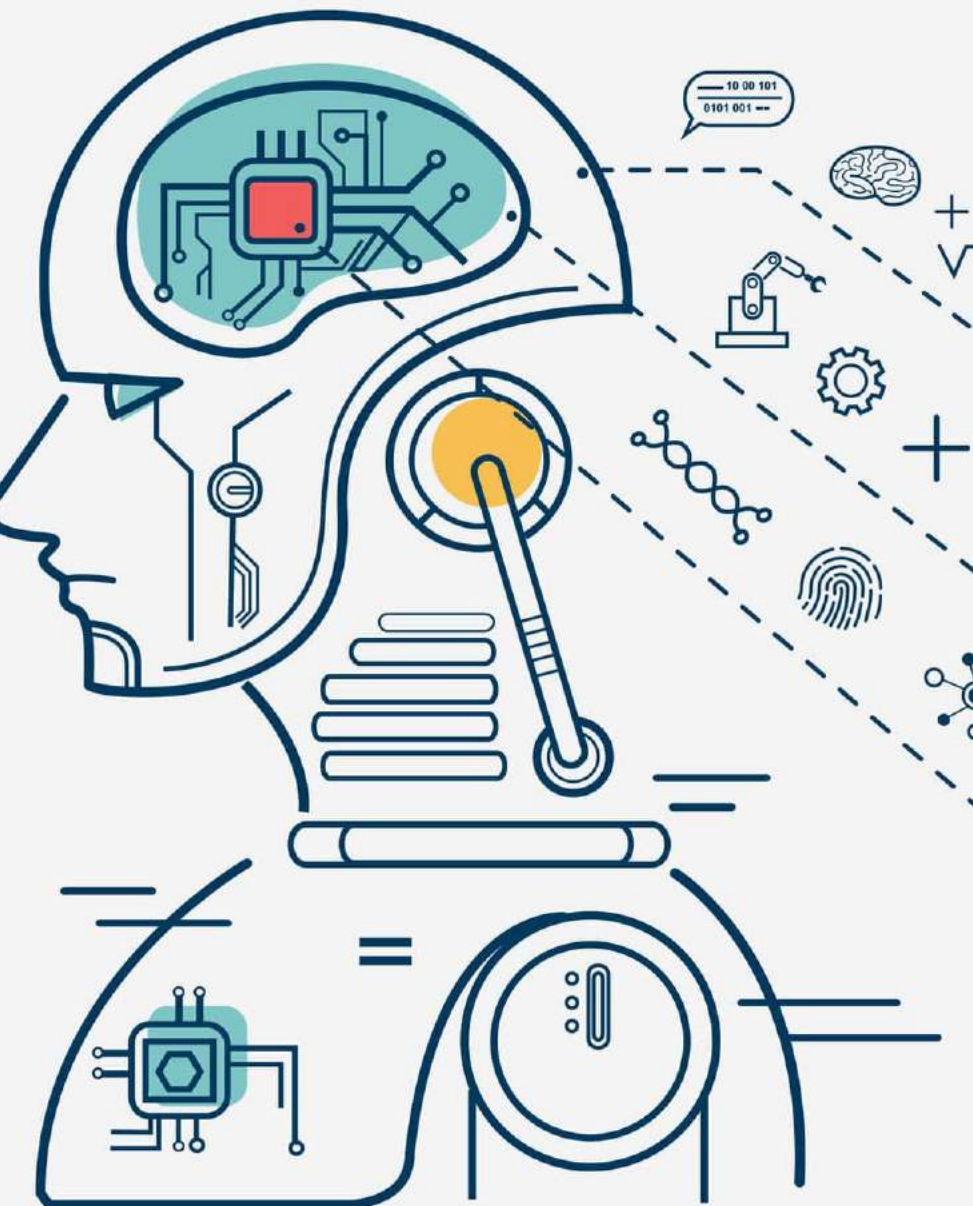


Feature Importance - Selection



- High Missing %
- Collinearity
- 0 importance
- Low importance
- Unique values
- Low variance





Goals for Talk

Feature Engineering

- Domain Expertise
- Automation tools

Feature Importance

- Calculation
- Selecting

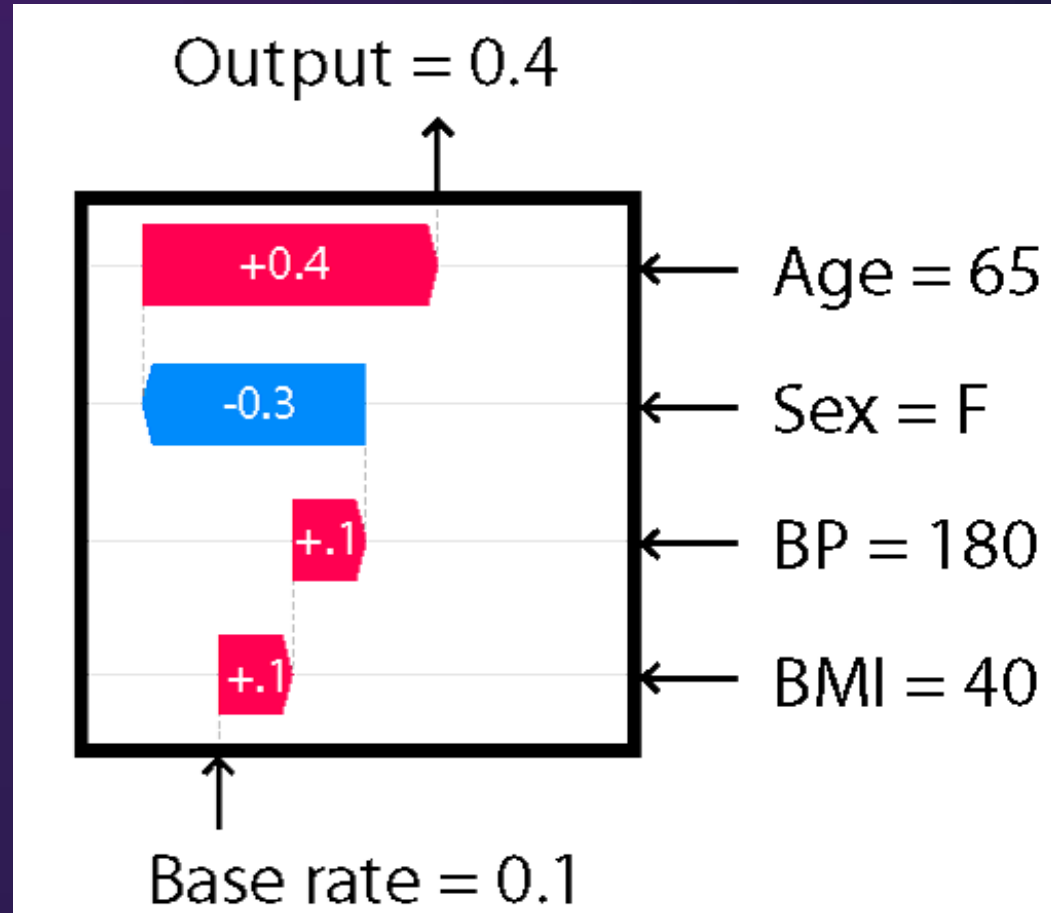
Modeling

- Interpreting
- Taking Action
- Knowing when to ignore the machine



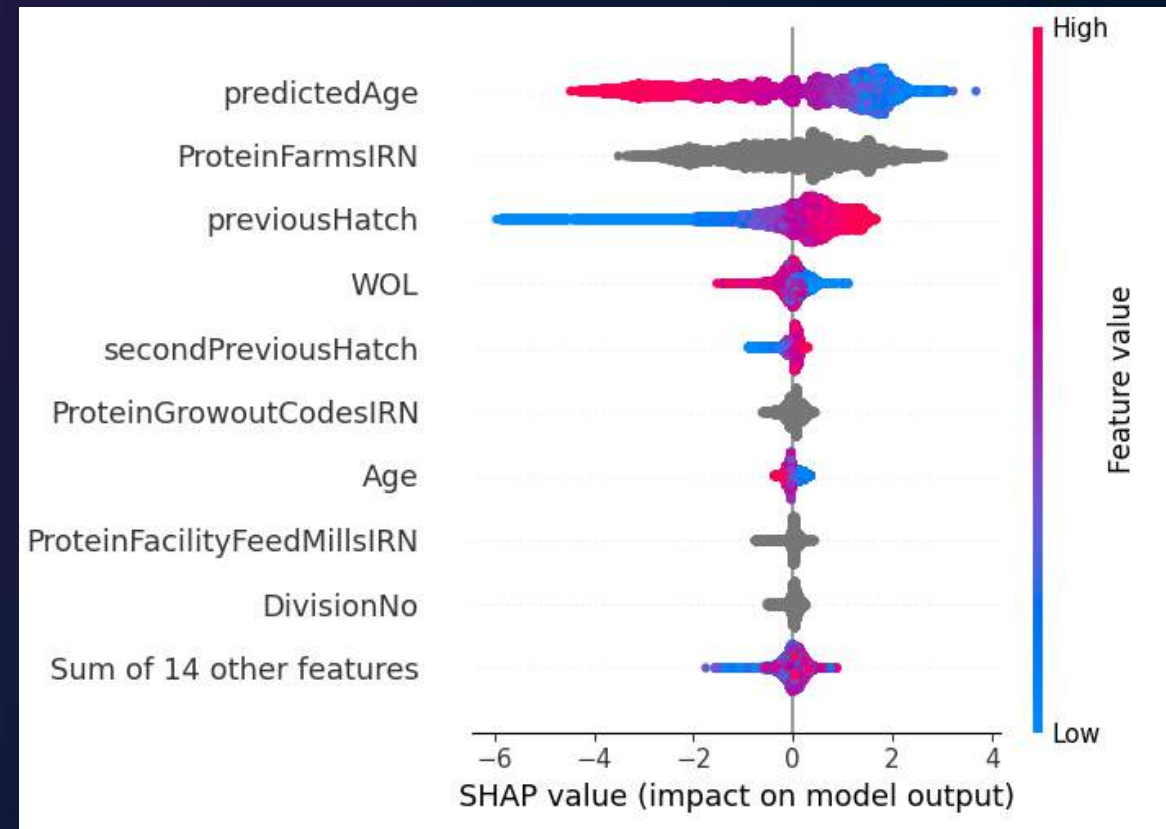
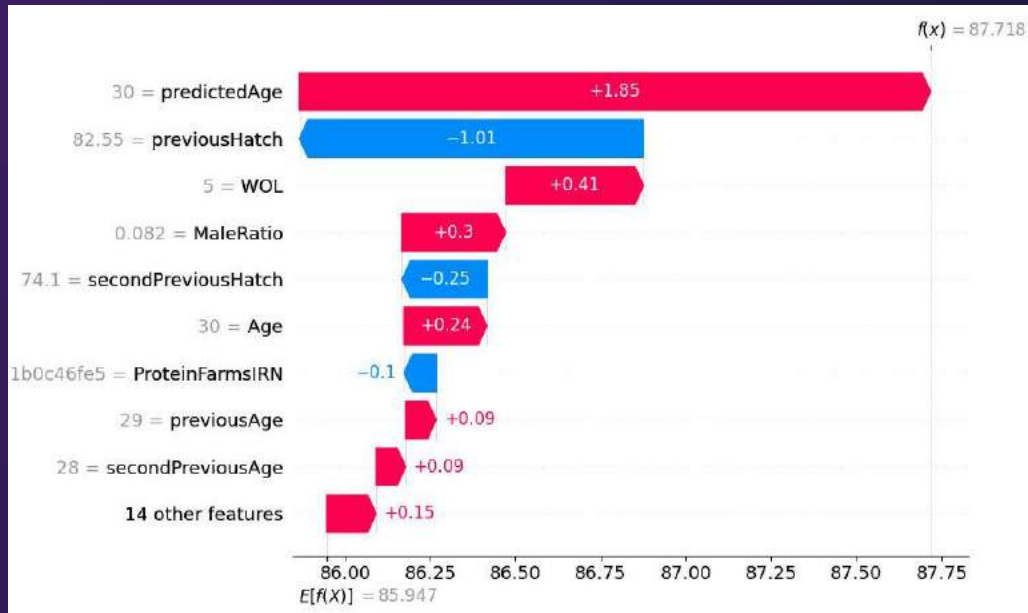
September 20th-22nd, 2022
MTech Users Conference

Modeling - Interpreting

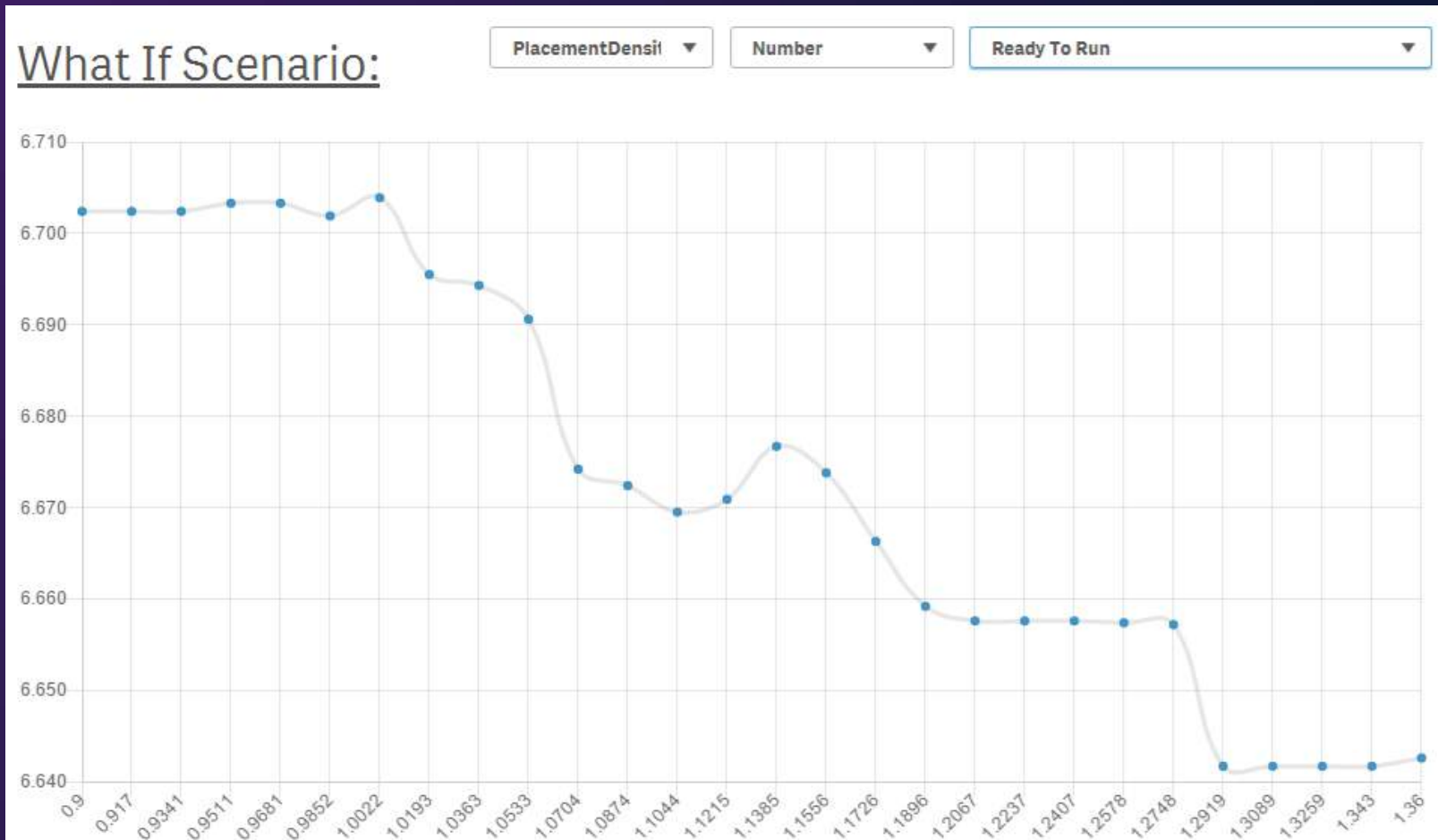




Modeling - Interpreting



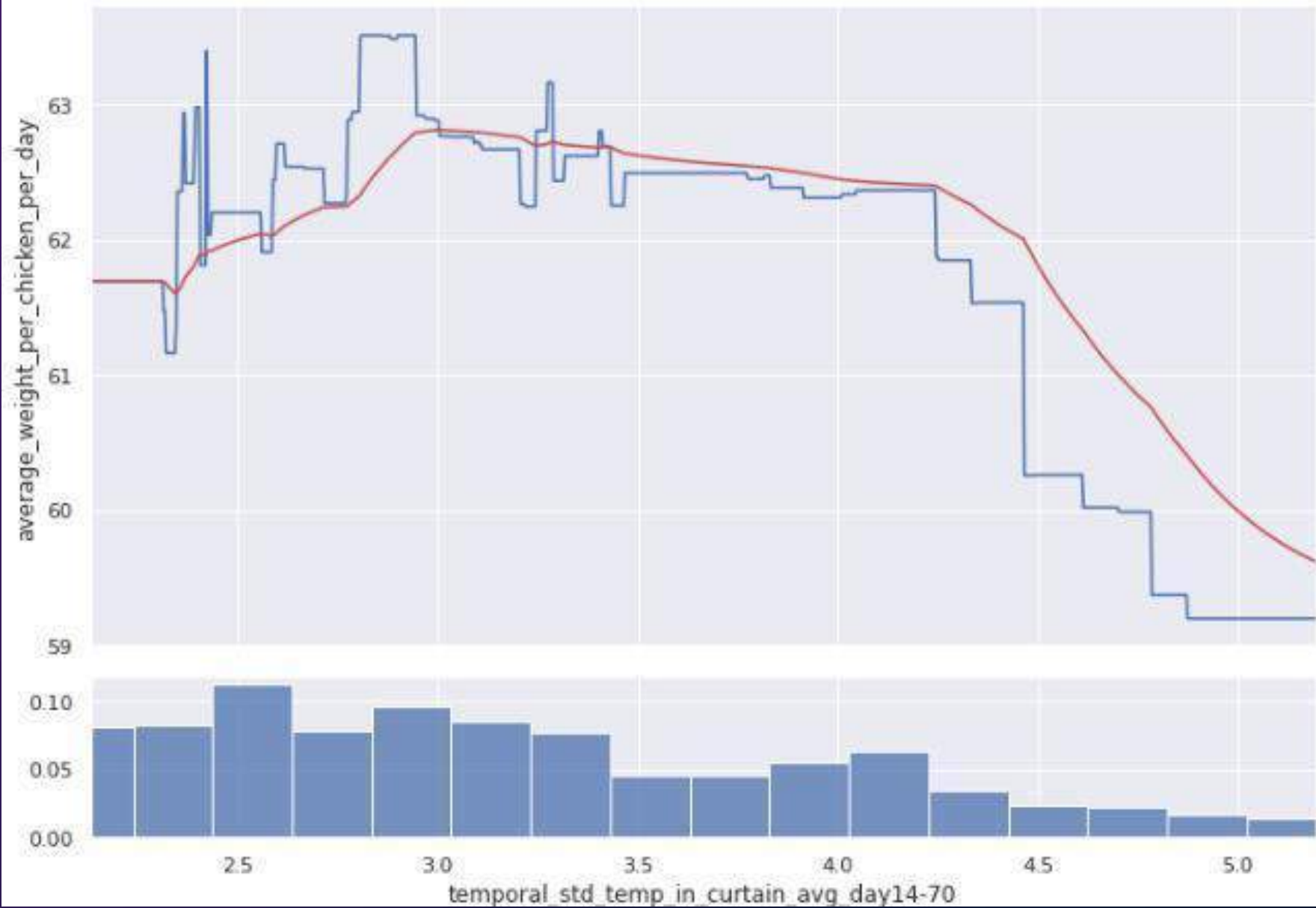
Modeling – Taking Action

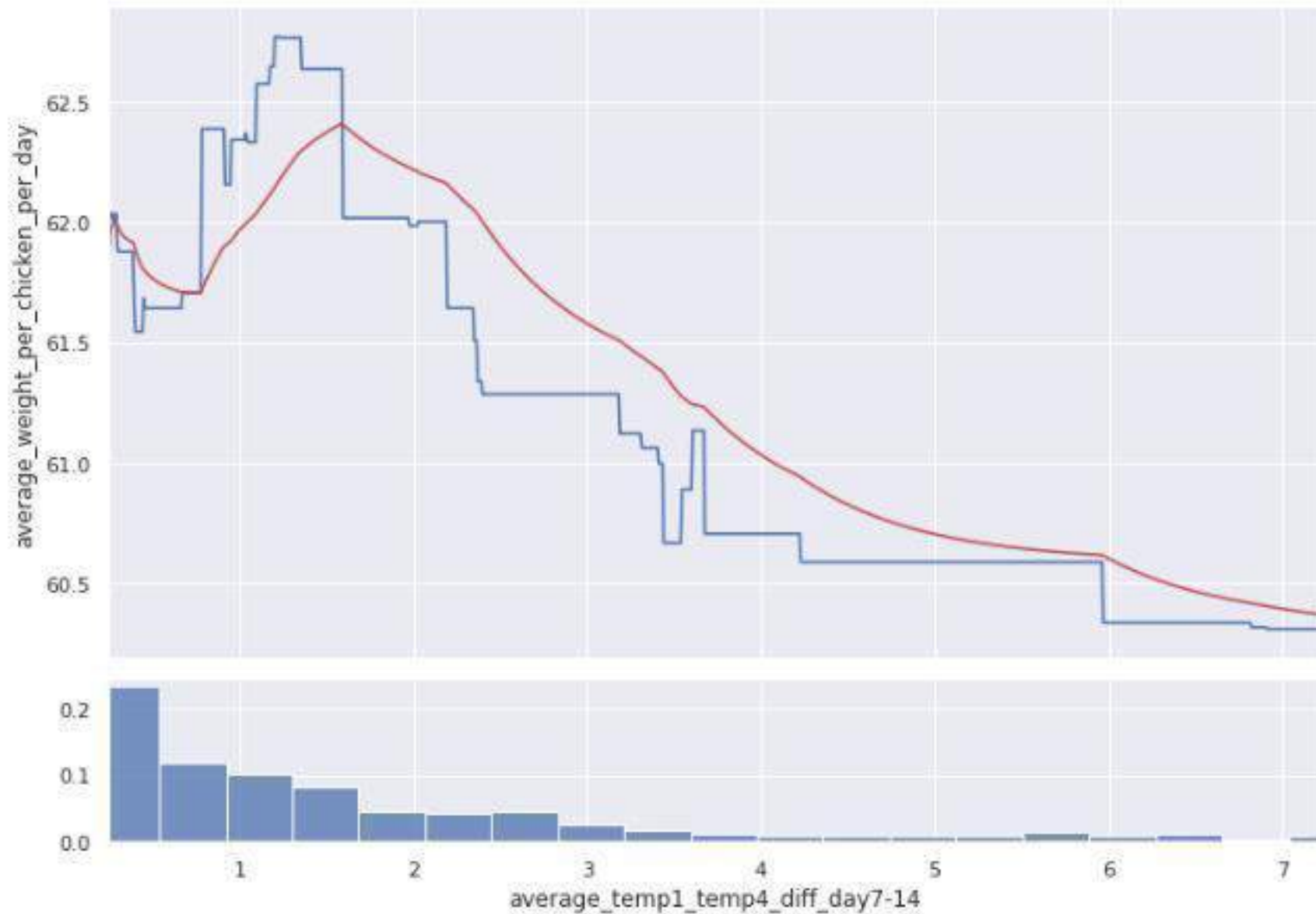


September 20th-22nd, 2022
MTech Users Conference

Modeling – When to Ignore the Machine









September 20th-22nd, 2022
MTech Users Conference

Thanks Gracias 謝謝 اركش Danke Спасибо ありがとう Obrigado

mtechsystems.io