Matching Yeast Polarity with Semiconductor Polarity and Interfacial Properties: A Route to Decode Electro-Genetic Information

Anna Grumman, Albena Ivanisevic[1]

[1] Department of Materials Science and Engineering, North Carolina State University, Raleigh, NC 27695

Abstract

This research uses a data-driven approach to explore how material properties can be altered to encode a cell response in yeast. The main focus of the project is organizing data that measures how accumulation of reactive oxygen species (ROS) causes cell death and triggers genetic changes in yeast. This involves extracting data from literature, storing and retrieving it in a relational database using SQL, and uploading it to Citrine Informatics' machine learning platform for materials data. These tools can be useful resources in predicting what properties are important in affecting cell pathways in yeast.
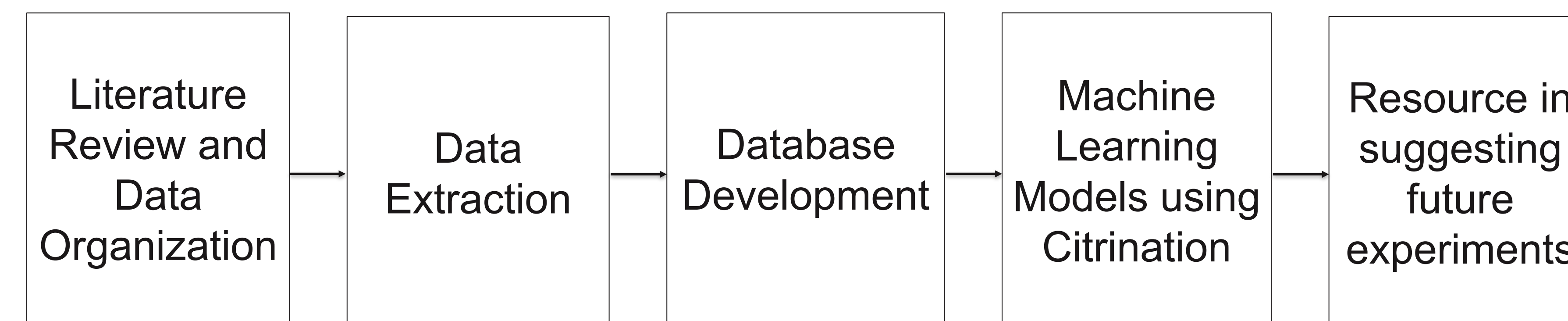
Background

- The properties of a material can be altered to achieve certain biological affects in yeast
- For example, persistent photoconductivity (PPC) can be induced in GaN, causing a buildup of reactive oxygen species which can ultimately induce cellular responses in yeast such as a reduction in MMP[1]
- Yeast cells have defensive mechanisms to protect from small doses of reactive oxygen, but in high quantities, ROS can cause cell death and DNA damage[2]
- Understanding how reactive oxygen species can trigger genetic changes in yeast can help determine how the surface chemistry should be altered

Excess carrier accumulation after backside UV illumination
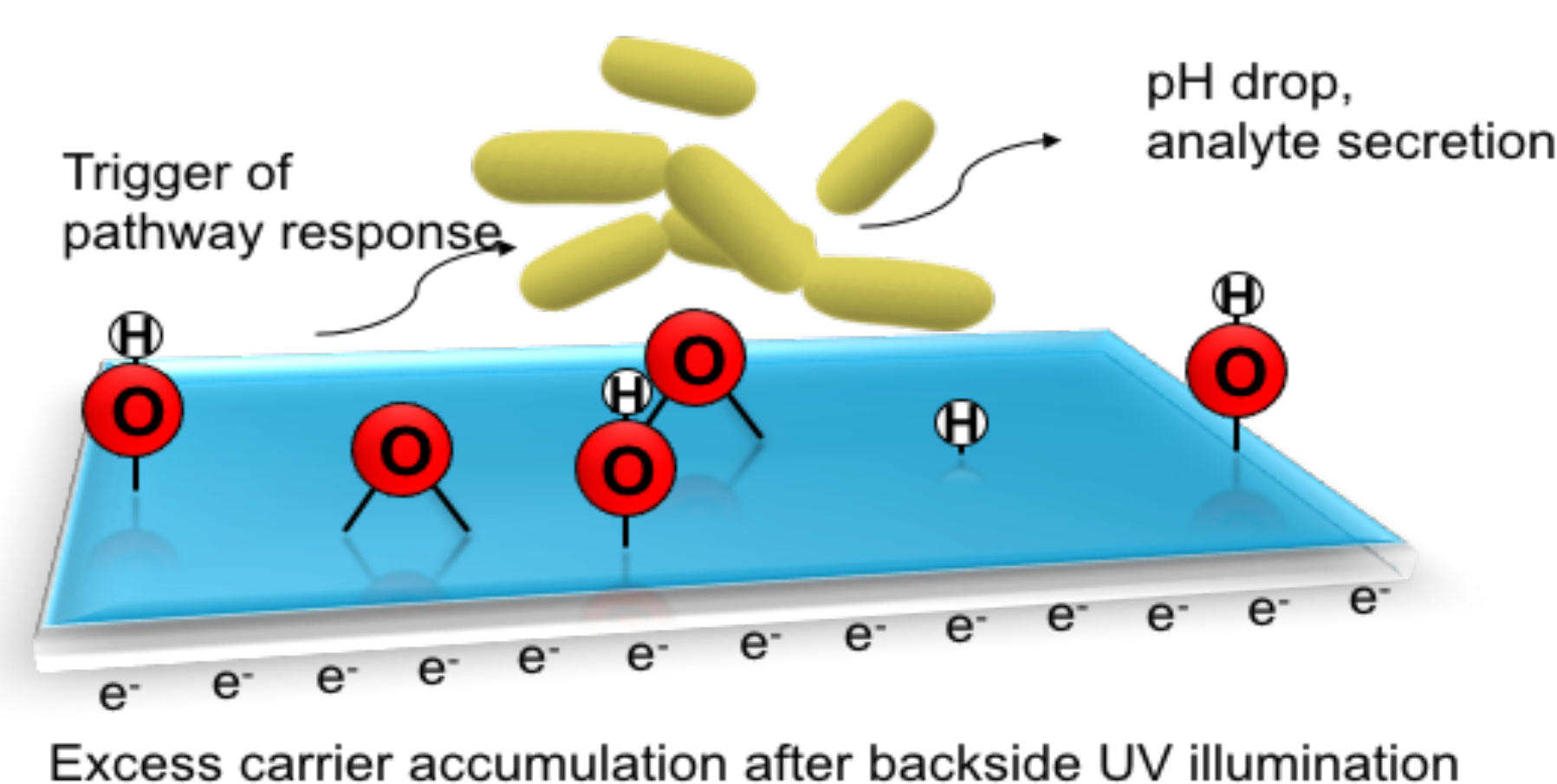
Methods

- Data measuring reactive oxygen species affect yeast cell pathways exists but is disorganized and not uniformly quantified
- The main focus of this project was a literature review determining what data is important and how it is measured
- Data was extracted from literature and stored in a database using SQL
- The data could also be uploaded to Citrine Informatics' machine learning platform, Citrination, to generate predictive models
- Ultimately, these tools could be resources in predicting what material properties affect cell pathways and suggest possible future experiments

Literature Review and Data Organization → Data Extraction → Database Development → Machine Learning Models using Citrination → Resource in suggesting future experiments

Data Organization

- Extracted published data from graphs using Citrine Informatics' Digitizer tool and from tables using Tabula
- Primary datasets measured concentration of reactive oxygen species and percent survival of yeast cells over time
- Other variables include genetic changes to yeast cells, type of reactive oxygen species, pretreatment of cells, temperature, and medium
- Stored data in a relational database using SQL for easy retrieval and analysis
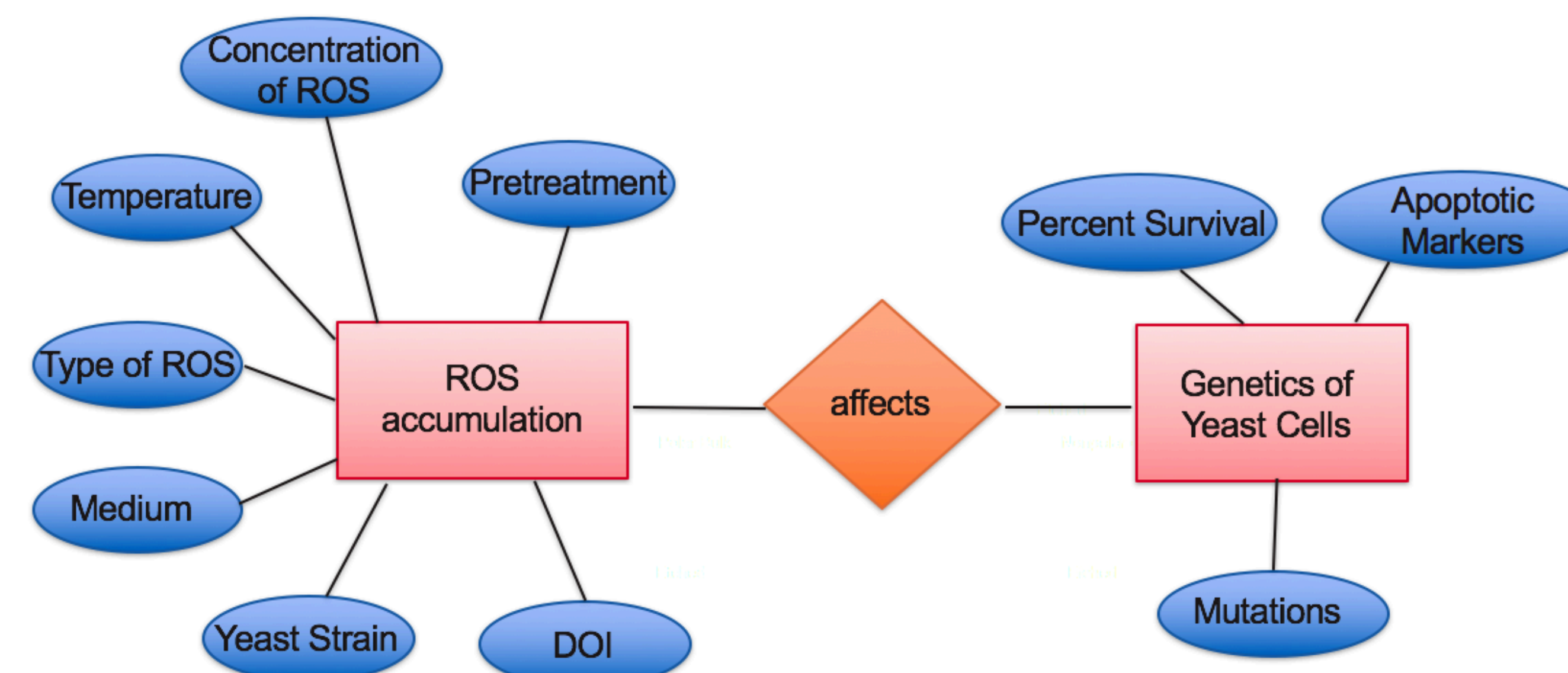
ER Diagram example for this research problem

Citrination

- Trained a machine learning model on time as an input and percent survival of yeast cells as an output
  - Removing the outliers from the dataset lowered the non-dimensional error below 1
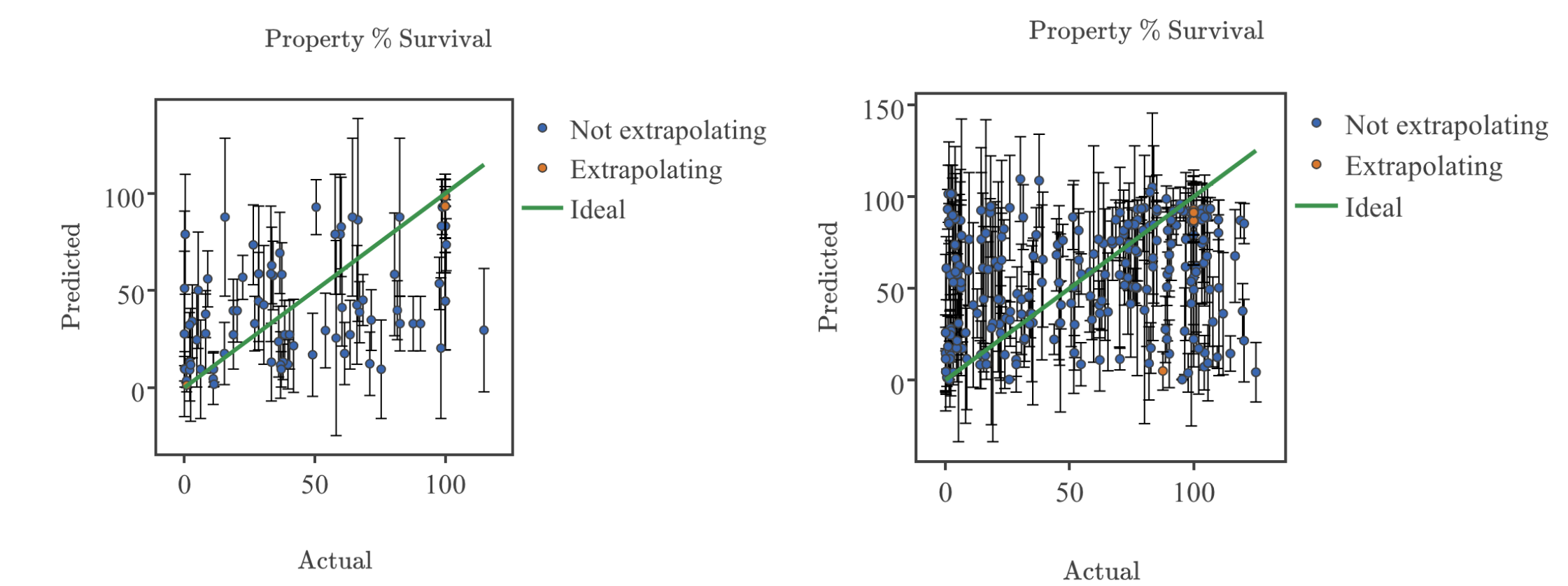
Predicted vs. Actual graphs with (left) and without (right) outliers of the model plotted against the actual values

- Working to build a model including all the inputs (real and categorical) as shown in the table to the right

Column Name	Descriptor Type	Parameter Type
Name	Categorical	Input
% Survival	Real	Output
Time	Real	Input
ROS	Categorical	Input
Concentration of ROS	Real	Input
Medium	Categorical	Input
Pretreatment	Categorical	Input
Genetic Changes	Categorical	Input
Temperature	Real	Input

- Also trained a model on concentration of ROS as an input and percent survival of yeast cells as an output with a different dataset
  - Non-dimensional error was 0.84

Future Work

- Working to automatically update data in the SQL database when a connected Excel spreadsheet is updated using SQLServer
- More data is needed to improve the current models to be able to predict what material properties trigger genetic changes in yeast

Funding & Acknowledgements

References

[1] Snyder, P. J.; Lajeunesse, D. R.; Reddy, P.; Kirste, R.; Collazo, R.; Ivanisevic, A. Nanoscale 2018, 10(24), 11506–11516.
[2] Perrone, G. G.; Tan, S.-X.; Dawes, I. W. Biochim. Biophys. Acta (BBA) – Mol. Cell Res. 2008, 1783(7), 1354–1368.