Caring with Algorithms

Jon Wayland

Director of Data Science
Florida Blue

Jonathan.Wayland@bcbsfl.com
https://www.linkedin.com/in/jonwayland

Machine Learning at Florida Blue

Care Management

Future Cost

Heart Transplants

Rewards and Incentives

Provider Insights

Preventable ER Visits

Product Support

Retention
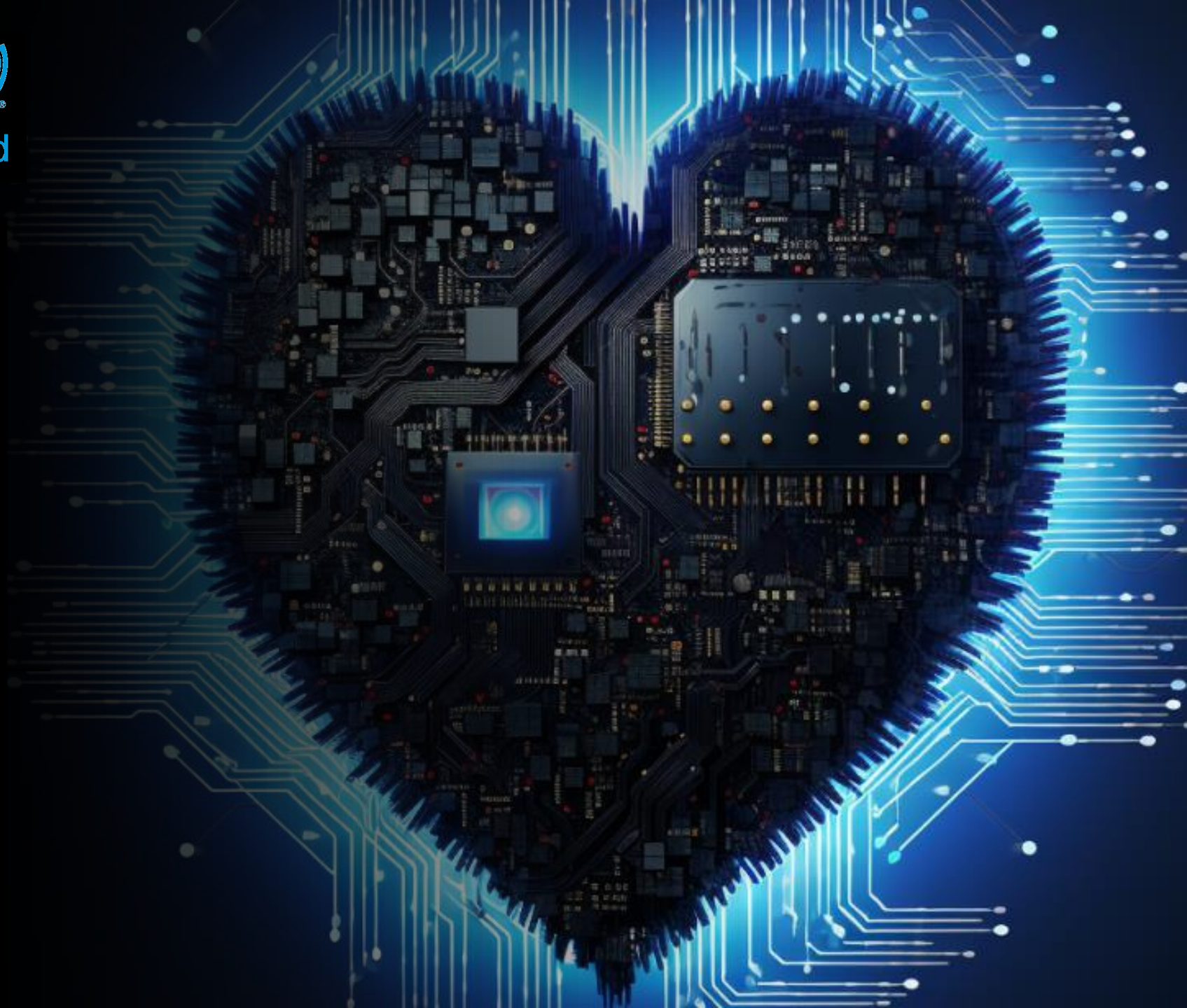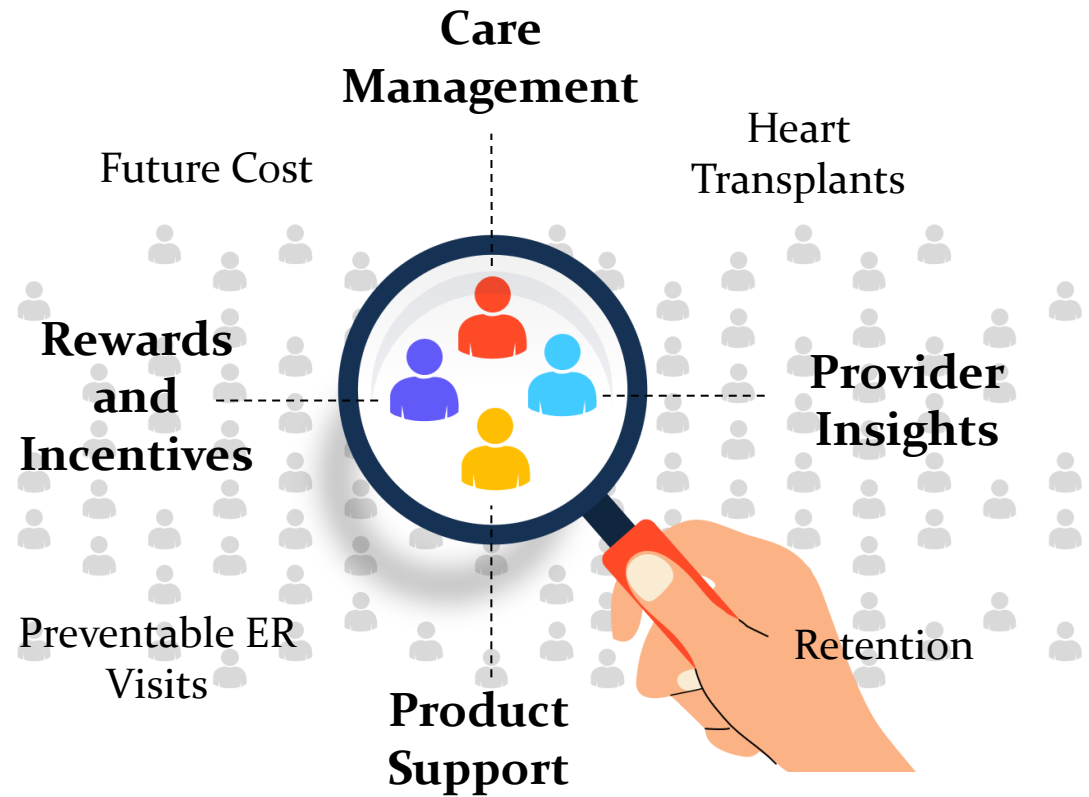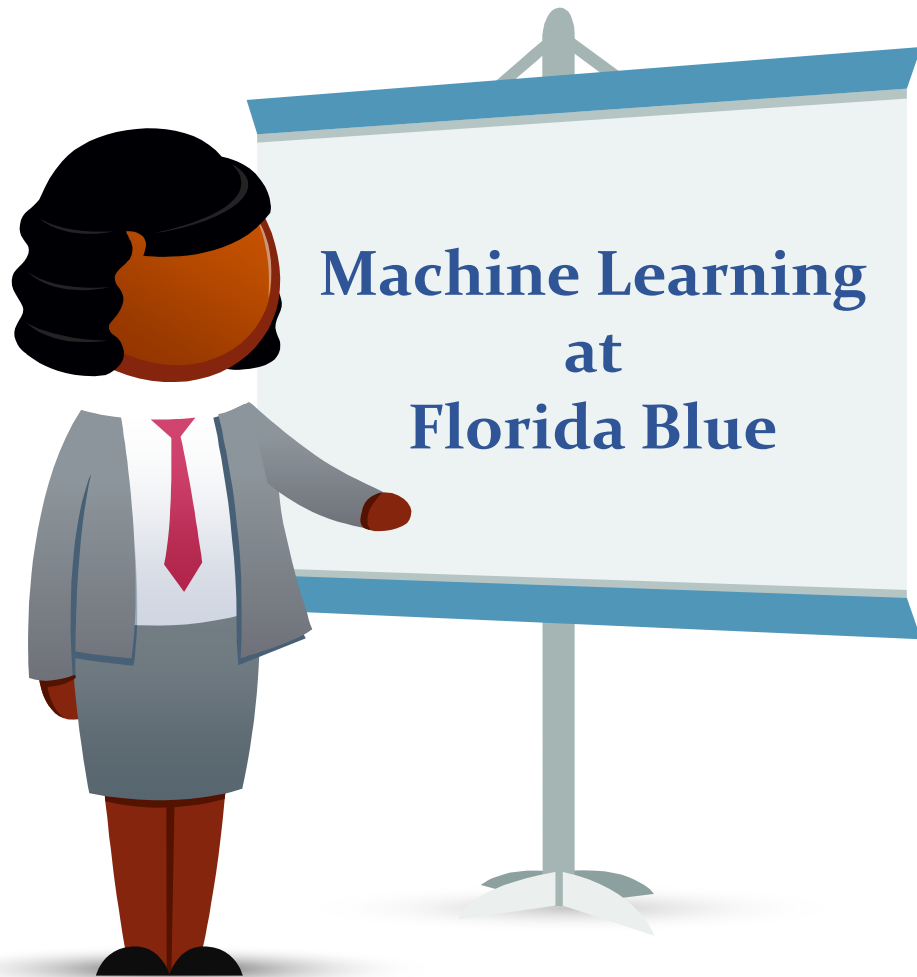
Care Continuum

Wellness ----------- Healthy members, no immediate needs for care

Acute ----------- Something is going on, but care is not urgent

At-Risk ----------- Care is needed to prevent ongoing problems

Chronic ----------- Continued care is required

Complex ----------- Multidisciplinary care is needed

End of Life ----------- Palliative care is coming soon

Cost & Risk Increases

Wellness → EOL

Membership Decreases

Shortness of Breath

Wellbeing

Depression

Pain

Nausea

Anxiety

Drowsiness

Appetite

Tiredness

Advanced Care Planning

PER MEMBER PER MONTH COST

Hospital
Hospice

MONTHS BEFORE DATE OF EXPIRATION

12 11 10 9 8 7 6 5 4 3 2 1 0
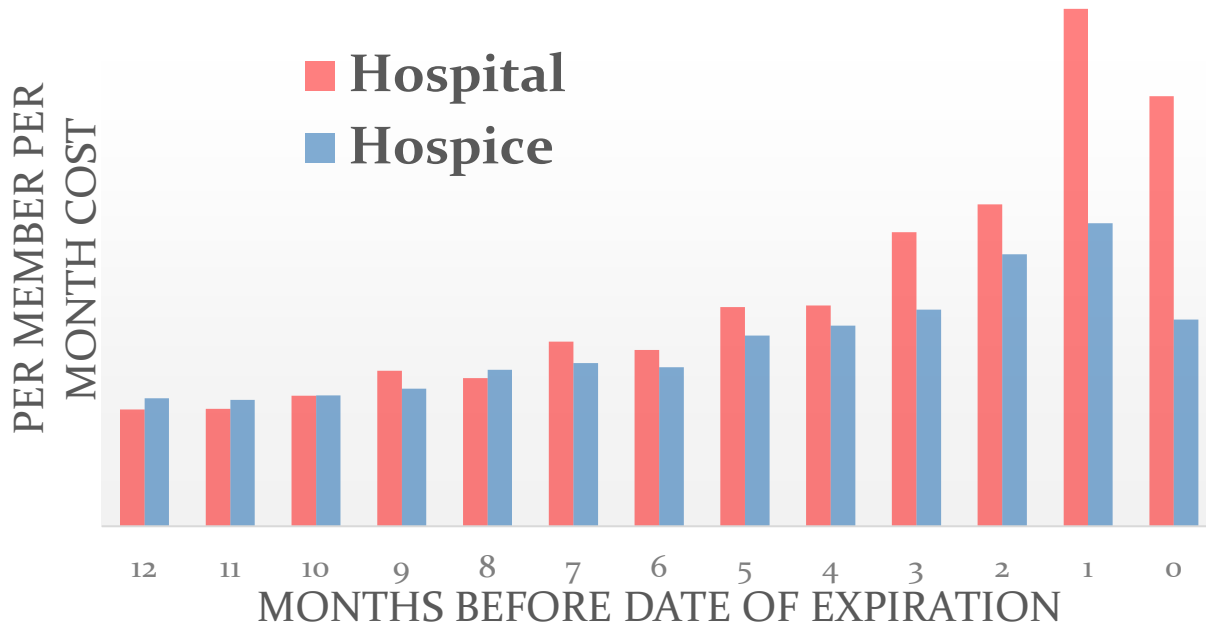
- Costs are 3x more in the final year of life than in the year prior

- EOL costs are 46x higher than the rest of the population

❑ Market need identified for Advanced Care

❑ Hospice vs. hospital deaths proxies financial opportunity

❑ Cross-functional solution designed
  ✓ Identification
  ✓ Advanced Directives
    o *Vendor offers identification in bundle*
  ✓ Palliative Care Delivery

Unsexy Part of Machine Learning

Building the training data

Removing Acute Events

- Clinical Input
- Research
- Cause of Death Analysis

Identifying Mortality

- Reported death dates (historically underreported)
- Vendor return data
- Discharge status codes

Code	Description
20	Expired -used only when the patient dies
40	Expired at home (Hospice claims only) used only on Medicare and TRICARE claims for hospice care
41	Expired in a medical facility (hospital, SNF, Intermediate Care Facility, or free standing hospice) for hospice use only
42	Expired - place unknown -this is used only on Medicare and TRICARE claims for Hospice only

Myocardial Infarction	Glioblastoma	Ovarian cancer
Chronic Heart Failure	Pancreatic cancer (all types)	Non lymphoma lymph node cancer
Peripheral Vascular Disease	Mesothelioma	Cancer of the head, face, and neck
Dementia	Transmissible spongiform encephalopathies	Heart cancer
Chronic Obstructive Pulmonary Disease	Lung cancer (all types)	Fibrodysplasia Ossificans Progressiva
Peptic Ulcer Disease	Cirrhosis	Gallbladder cancer
Mild Liver Disease	End Stage Renal Disease	Esophageal cancer
Moderate to Severe Liver Disease	Liver and bile duct cancer (Cholangiocarcinoma)	Cardiac Arrest (Previous)
Diabetes	Acute respiratory failure with hypoxia	Leukemia, acute myelomonocytic
HIV / AIDS	Ischemic Heart Disease	Uterine cancer
Lymphoma	Pulmonary fibrosis	Stomach cancer
Leukemia	Amyotrophic lateral sclerosis (ALS)	Hypopharynx cancer
Diffuse intrinsic pontine glioma (brain stem)	Brain cancer	Ventricular arrhythmias

Unsexy Part of Machine Learning

Building the training data (Continued)

1 Year Window

Prediction Month → Reference Month

0-36 Month Feature Collection

0.2% Prevalence

Structuring the Outcome
- Event = 1 if member passes away
- Event = 0 if member survives

Reference Month: Month of Death for those who passed away and Randomly Selected Month for those who survived

Example:

ID	Status	Month w/ Qualifying Condition
123	Survived	201702
123	Survived	201703
123	Survived	201704
123	Survived	201705
321	Survived	201610
321	Survived	201611
321	Survived	201612

ID	Reference Month
123	201803
321	201709

Note: 1 year added to the randomly selected month w/ potentially terminal condition to represent the prediction window

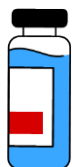# Feature Engineering
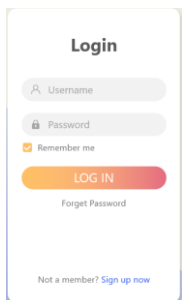
Enrollment

Contract Size
Age
Gender
Tenure w/ Florida Blue
Product Selection

SDoH

Area Deprivation (Block Level)
Social Vulnerability (Zip Code)
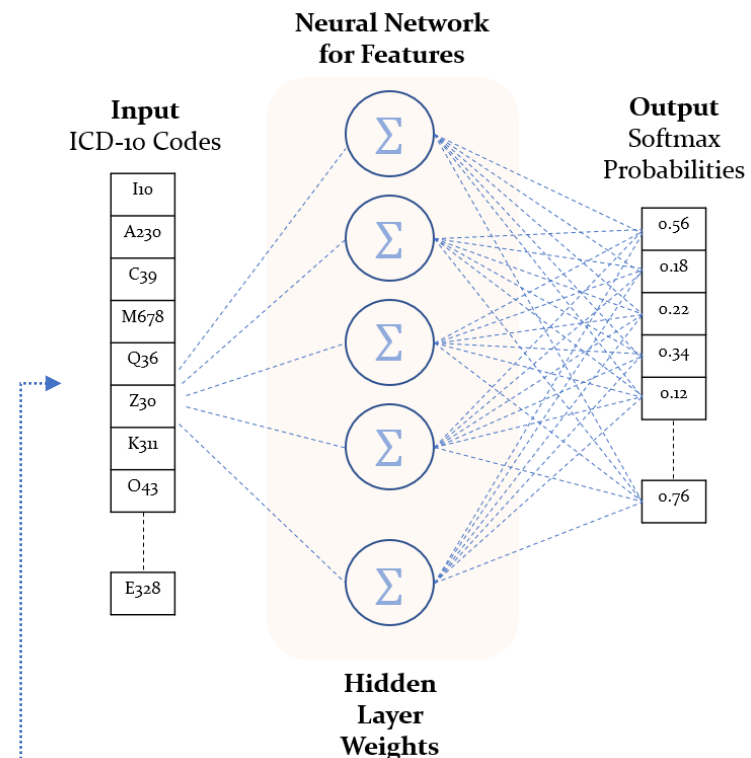CDC Local (Zip Code)

Pharmacy Claims

Script Counts
Script Costs
Utilization by Therapeutic Classes

Medical Claims

Claim Counts
Costs
Procedures
Diagnoses
Chronic & Disabling Conditions
High Death Rate Diseases
Cancer
Utilization by Provider Specialty
Charlson Comorbidity Index
Alcohol, Drug Dependence
Smoking
Obesity
Diagnosis Embeddings

Neural Network for Features

Input
ICD-10 Codes

I10
A230
C39
M678
Q36
Z30
K311
O43

E328

Output
Softmax Probabilities

0.56
0.18
0.22
0.34
0.12

0.76

Hidden Layer Weights

Note: Attention mechanism is superseding word2vec models

Final Data Preparation ✂

500 Features

Status	Age	Gender	...	Embedding 1	...
Survived	42	M	...	-0.234	...
Survived	46	F	...	0.563	...
Expired	55	M	...	1.891	...
...	...	...	...	...	...

99.8%

0.2%
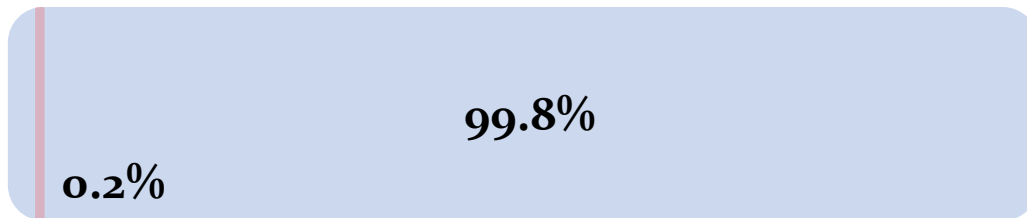
Filtering to Qualifying Condition Set

99%

1%
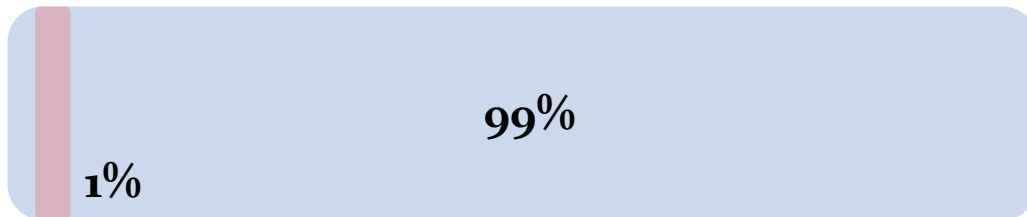
Training, Validation, Testing

70%   15%   15%

Over/Under Sampling

50%   50%
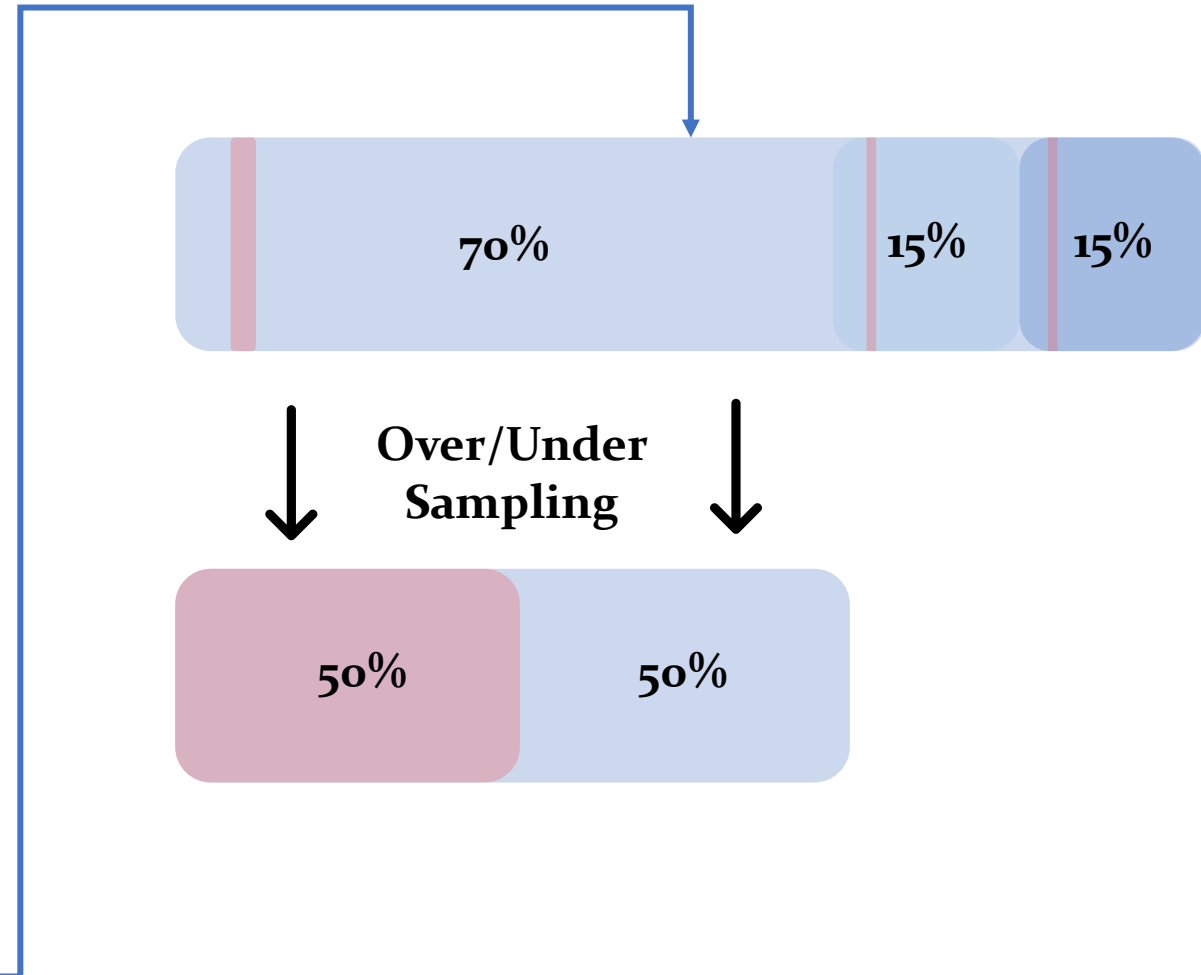
The Sexy Part of Machine Learning

Algorithm Selection
- Logistic Regression
- Penalized Logistic Regression
  - Ridge, LASSO, and Elastic Net
- Decision Tree
- Random Forest
- Support Vector Machines
- Naïve Bayes
- Gradient Boosting
- XGBoost

Validation Data

15%

Hyperparameter Tuning

Learning rate
Percent of variables to consider
Ratio of the training instances
Max depth of the tree
Number of iterations
Min loss reduction required for new partitions
Min sum of instance weight

Searching over 360 hyperparameter combinations

**10-Fold
Cross Validation**

**70% of Data
50/50 Prevalence**

*Repeated
Twice*

Performance on Testing Data

Precision on Top 0.2% = 35%
AUC = 0.90

Interpretations:

Random Selection of 1,000
2 Passing Away

Model Selection of 1,000
350 Passing Away

Expectation for Generalization

If both a surviving and expired member were randomly selected, then the probability that the model ranked the expired member higher than the surviving member is 90% (AUC)
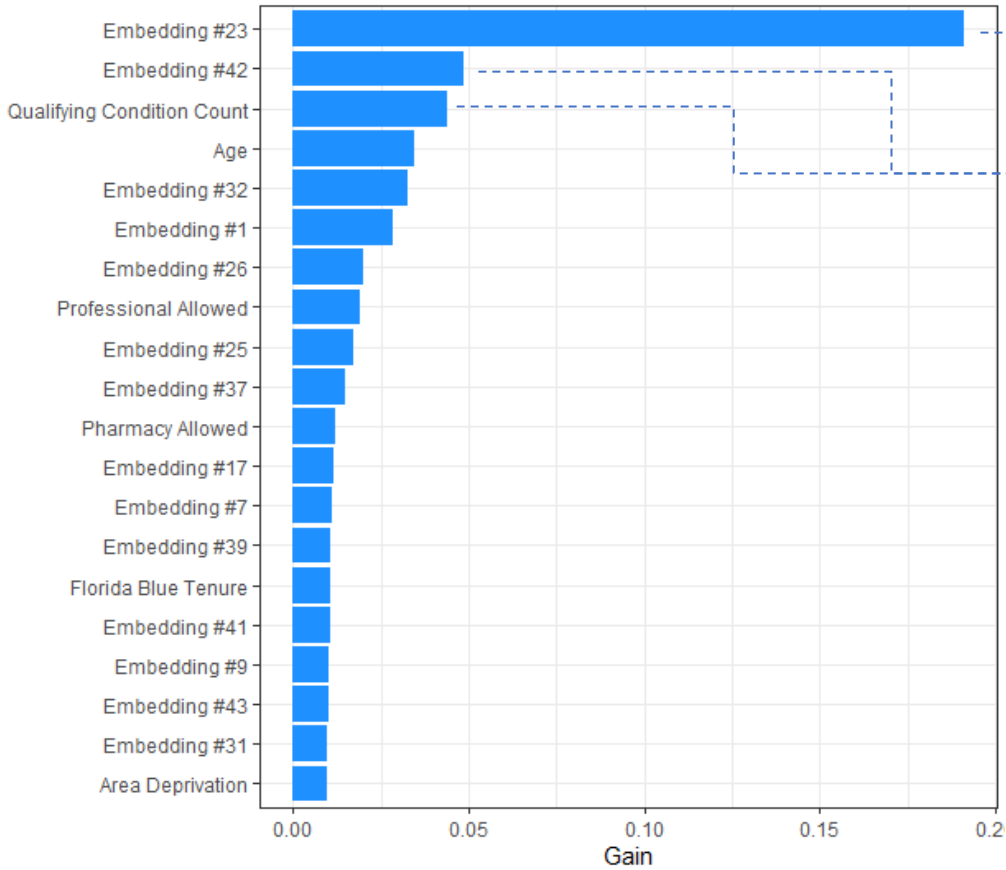
MidJourney bringing "XGBoost" to life

The XGBoost Model

Top 20 Features by Importance

Defending XGBoost

To Non-Technical Audiences

Decision Tree

Use When:
You want to understand impacts of variables on an outcome

✓ Highly Interpretable

✓ Mediocre Predictive Capability

Can easily understand every prediction by traveling back up the tree

Node 1

Node 2

Node 3

Node 4

No

Yes

Yes
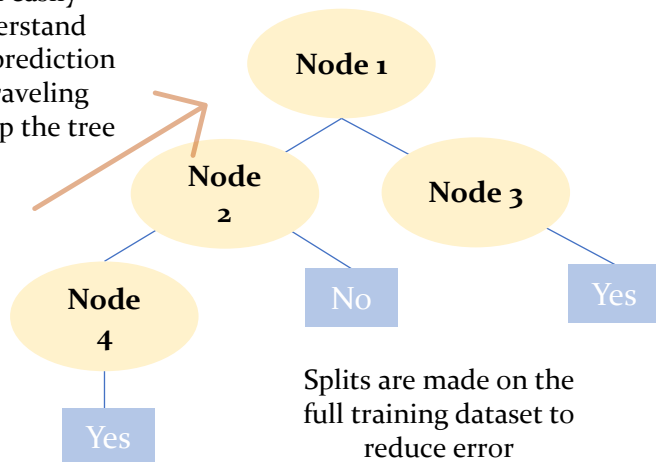
Splits are made on the full training dataset to reduce error

XGBoost

Use When:
You want to obtain the most accurate prediction possible

✓ Nominally Interpretable

✓ Strong Predictive Capability

Data Subset 1

Tree 1,1	Wrong	Right	Right	Wrong	Right
Tree 1,2	Right			Wrong	
Tree 1,3				Right	

Each tree focuses on the misclassified observations from the prior tree

Optimized to run as fast as possible given available resources

Data Subset 2

Tree 2,1	Right	Wrong	Right	Right	Wrong	Wrong
Tree 2,2		Wrong			Right	Wrong
Tree 2,3		Wrong				Right
Tree 2,4		Right				

Error is minimized using gradient descent

XGB Model
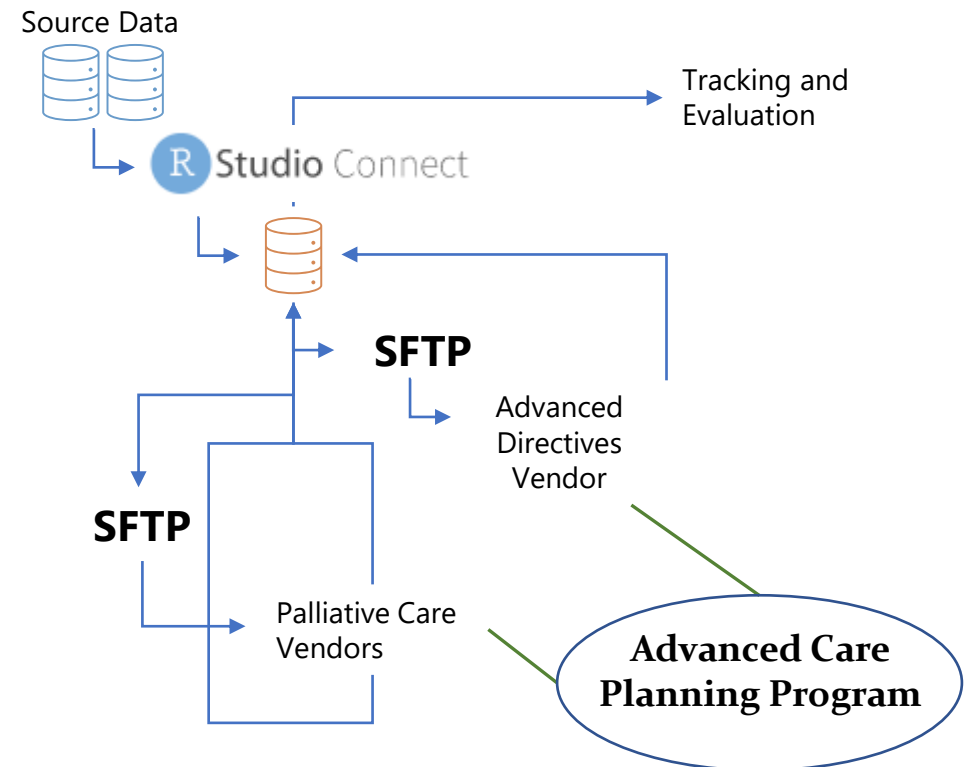
Predictions cannot be easily traced back

Flipping the "On" Switch

Business Value Machine Learning Framework
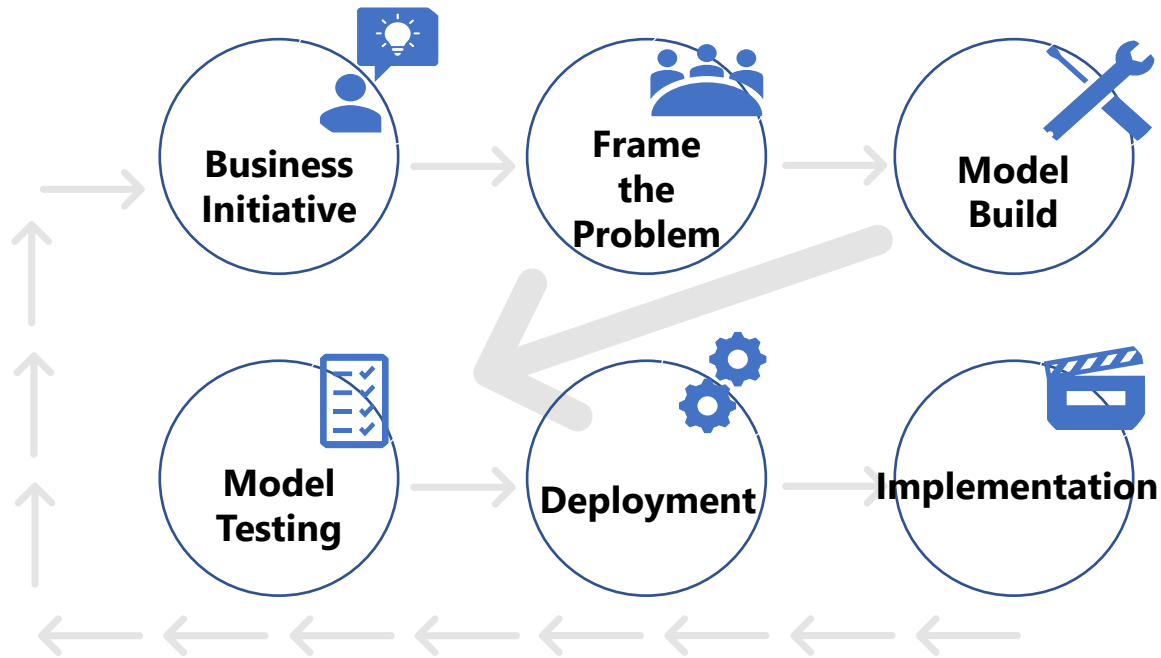
Deployment-Implementation Framework

So What?

Recall

- Financial Opportunity -

- Testing Precision: 35% -

- Vendor sells ML model for other clients - - - - - - - - - - - - - - - -

- These are real people -

After 2 Years of Experience

- Savings targets hit year-over-year

- 45% Precision

- Highest engagement rate in our vendor's book of business
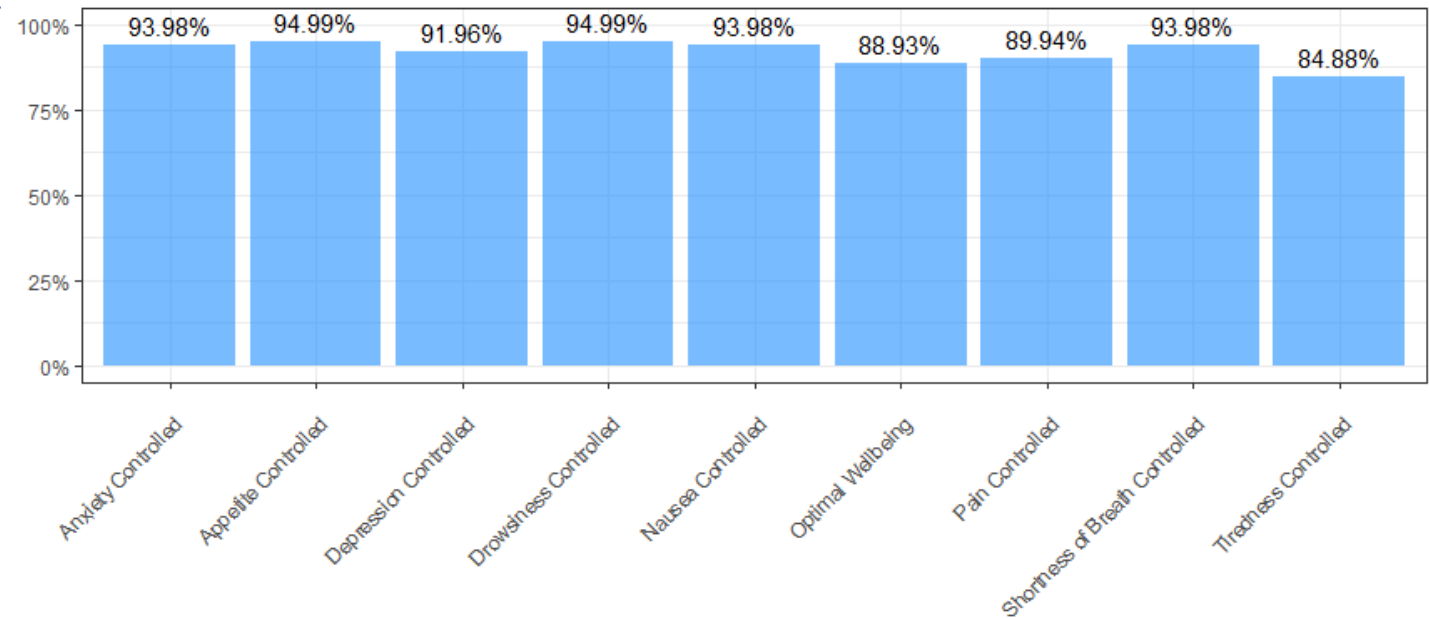
Jonathan.Wayland@bcbsfl.com
JonWayland47@gmail.com

https://jonwayland.com

Jon Wayland

Data Science Director @ Florida Blue

https://www.linkedin.com/in/jonwayland