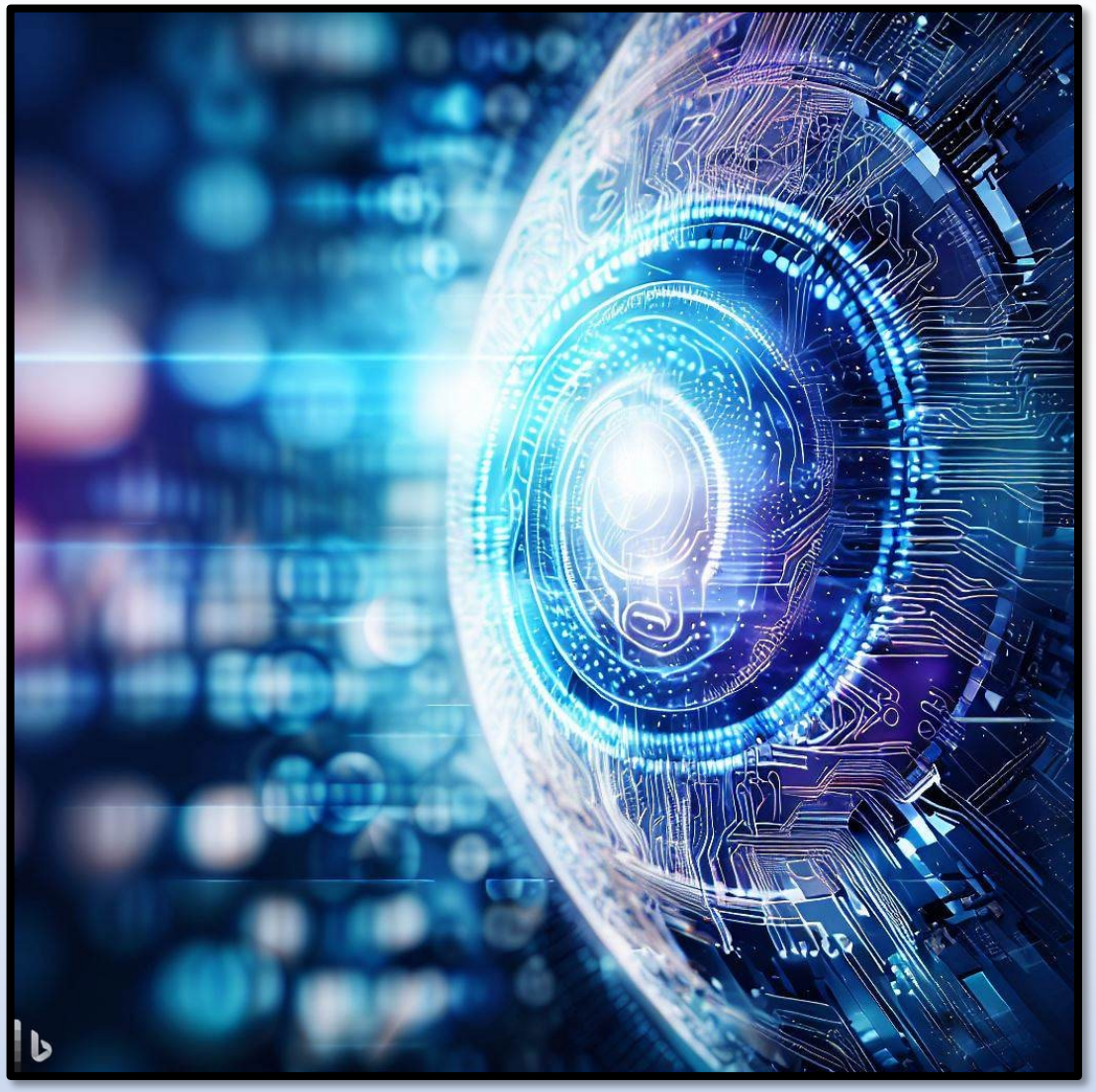
Unlocking New Insights in Epic Data with Predictive Analytics



David Patrishkoff MS LSSMBB
SUNY Polytechnic Institute

Robert E. Hoyt FACP FAMIA ABPM-CI
Virginia Commonwealth University

Sept 19, 2023
Quarterly Faculty Development Meeting
Mohawk Valley Health System

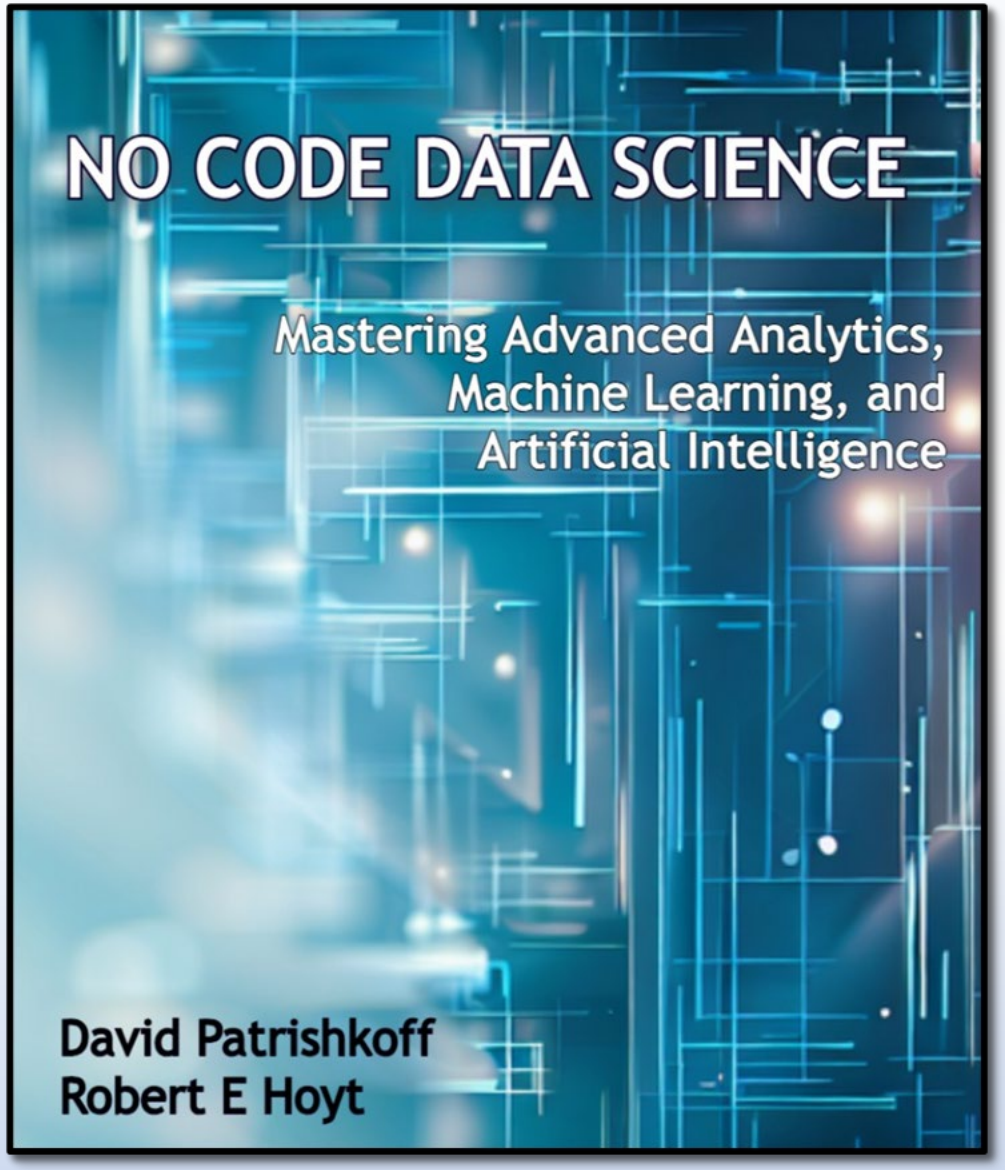
Agenda

1. Introduction
2. What is Data Science and Predictive Analytics?
3. Potential Epic Applications of ML/AI
4. No code Predictive Analytics of Epic data

Agenda

1. Introduction
2. What is Data Science and Predictive Analytics?
3. Potential Epic Applications of ML/AI
4. No code Predictive Analytics of Epic data

High-End No-Code Data Science (ML/AI) is Possible

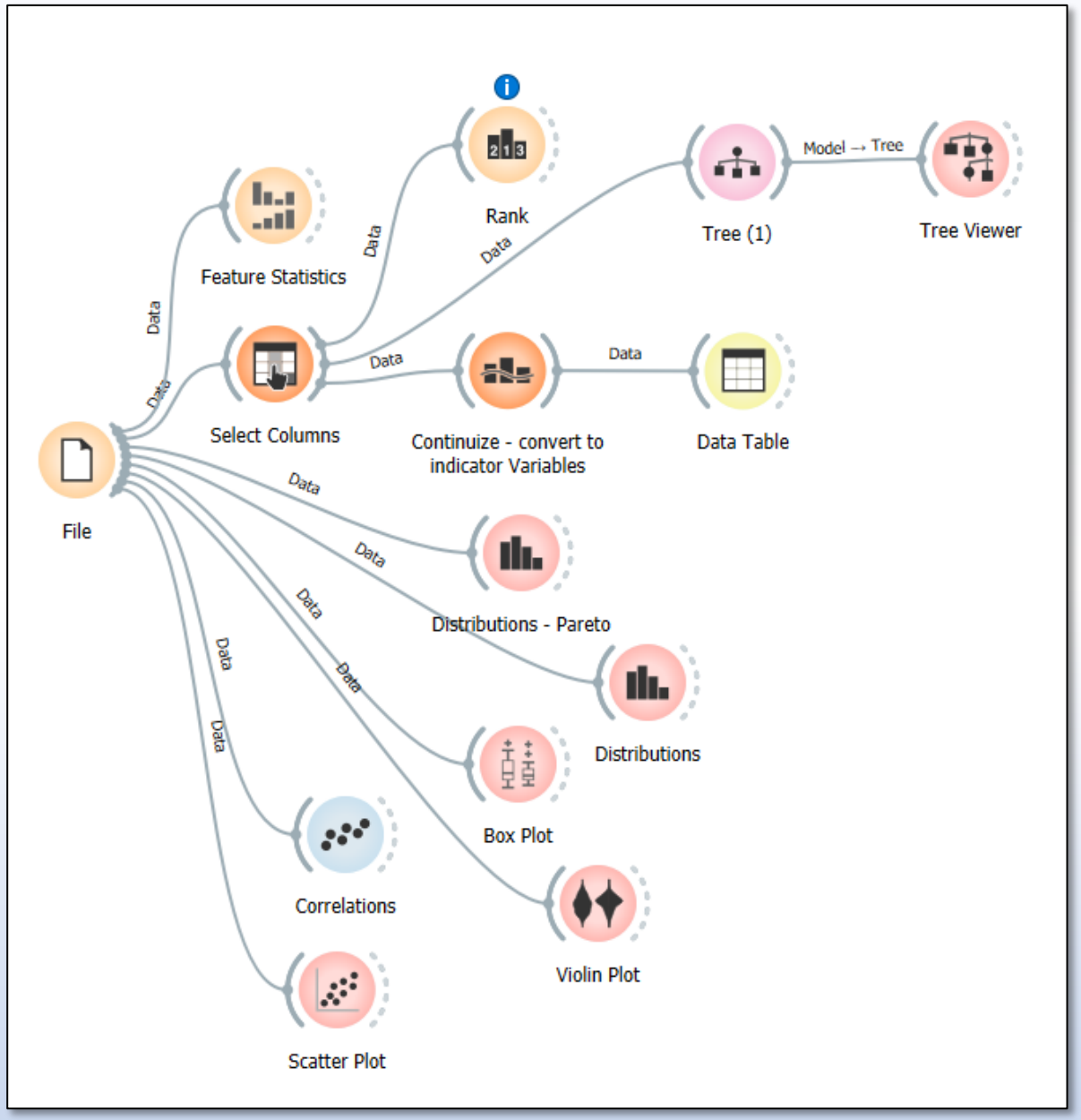


- No-Code open-source Software for ML to analyze tabular data, text mining, and image analysis, based on Python and R
- Predictive analytics with Orange, JASP and BlueSky Statistics Software
- Mayo Clinic Announces Move from SAS' JMP to BlueSky Statistics
- Multi-industry examples and applications
- Free monthly workshops ongoing
- Certification levels start in late 2023 / 2024
- We will link the ChattyPDF chatbot to our book to answer any question related to the book content

https://www.nocodedatascience.net/

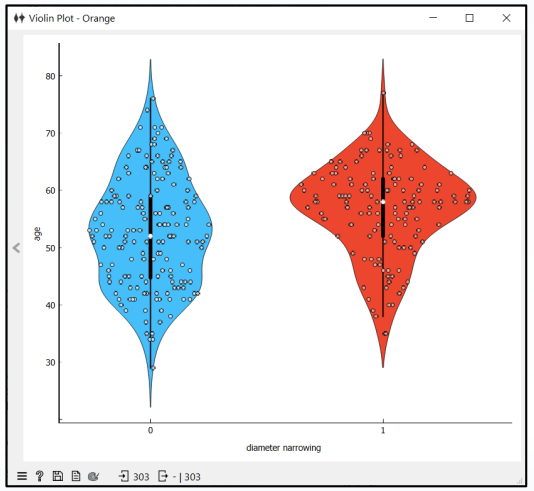
Manuscript completed: August 2023. Publishing target: Oct 2023.

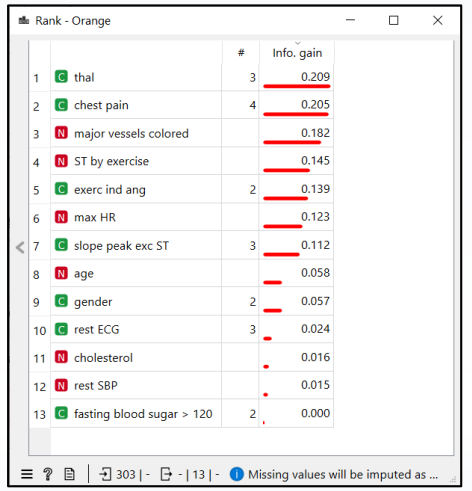
Orange Visual Programming for Predictive Analytics and AI

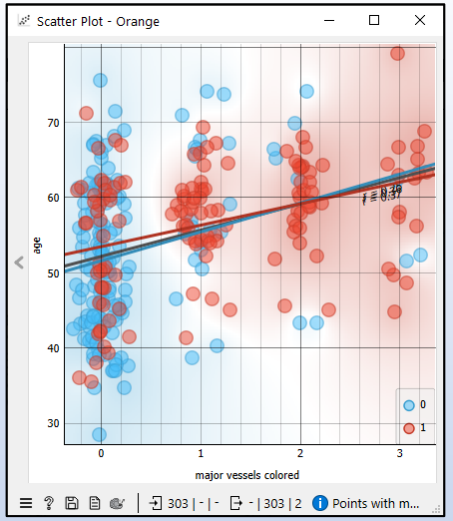


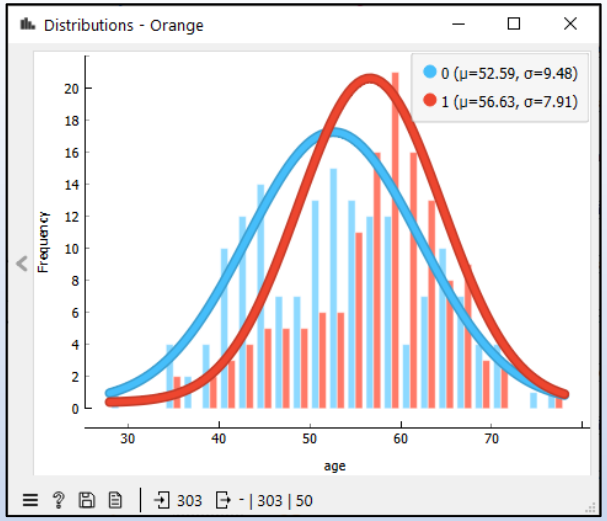
Exploratory Data Analysis:

This is a workflow that can be used for any dataset to explore

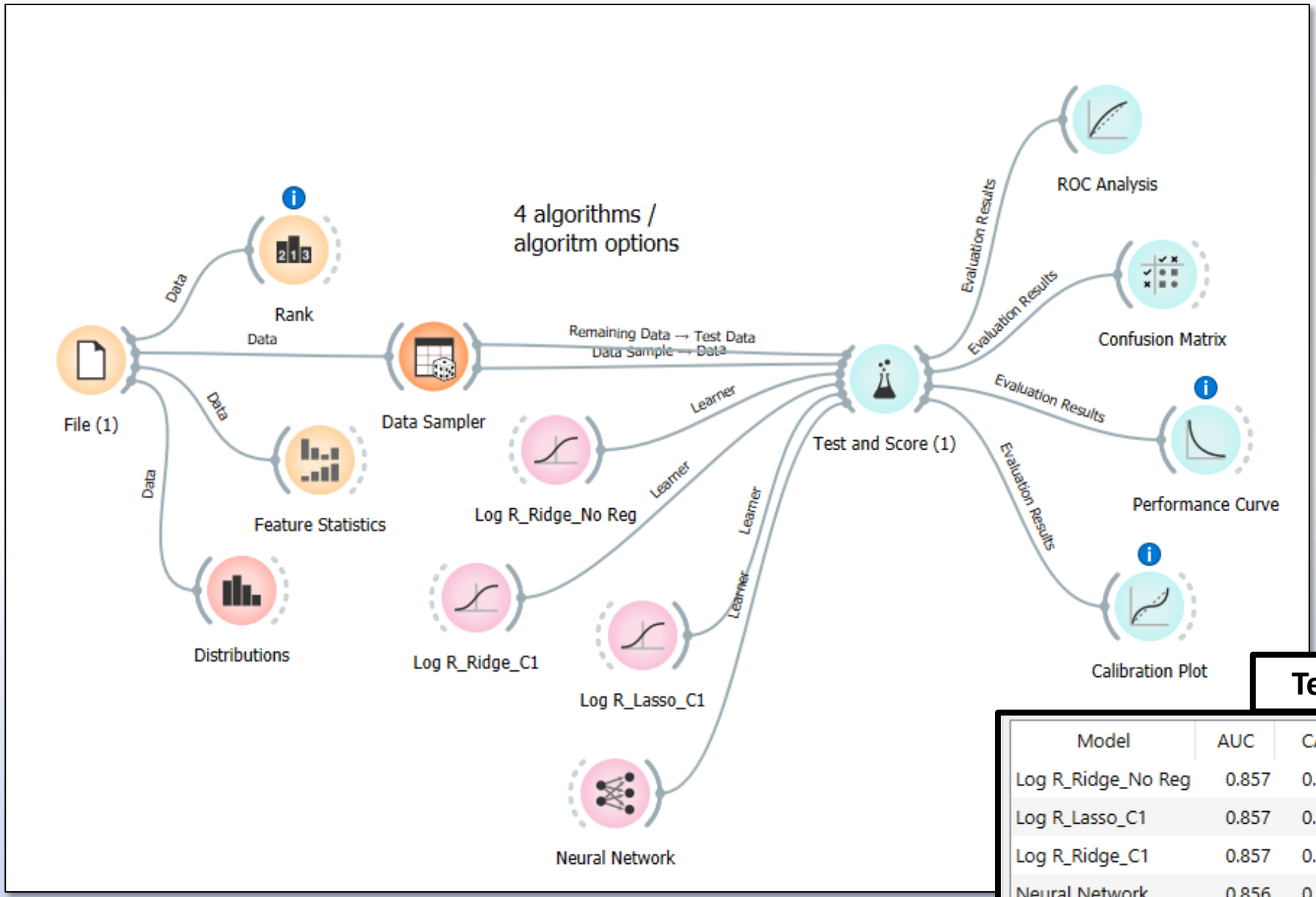








Orange Visual Programming for Predictive Analytics and AI



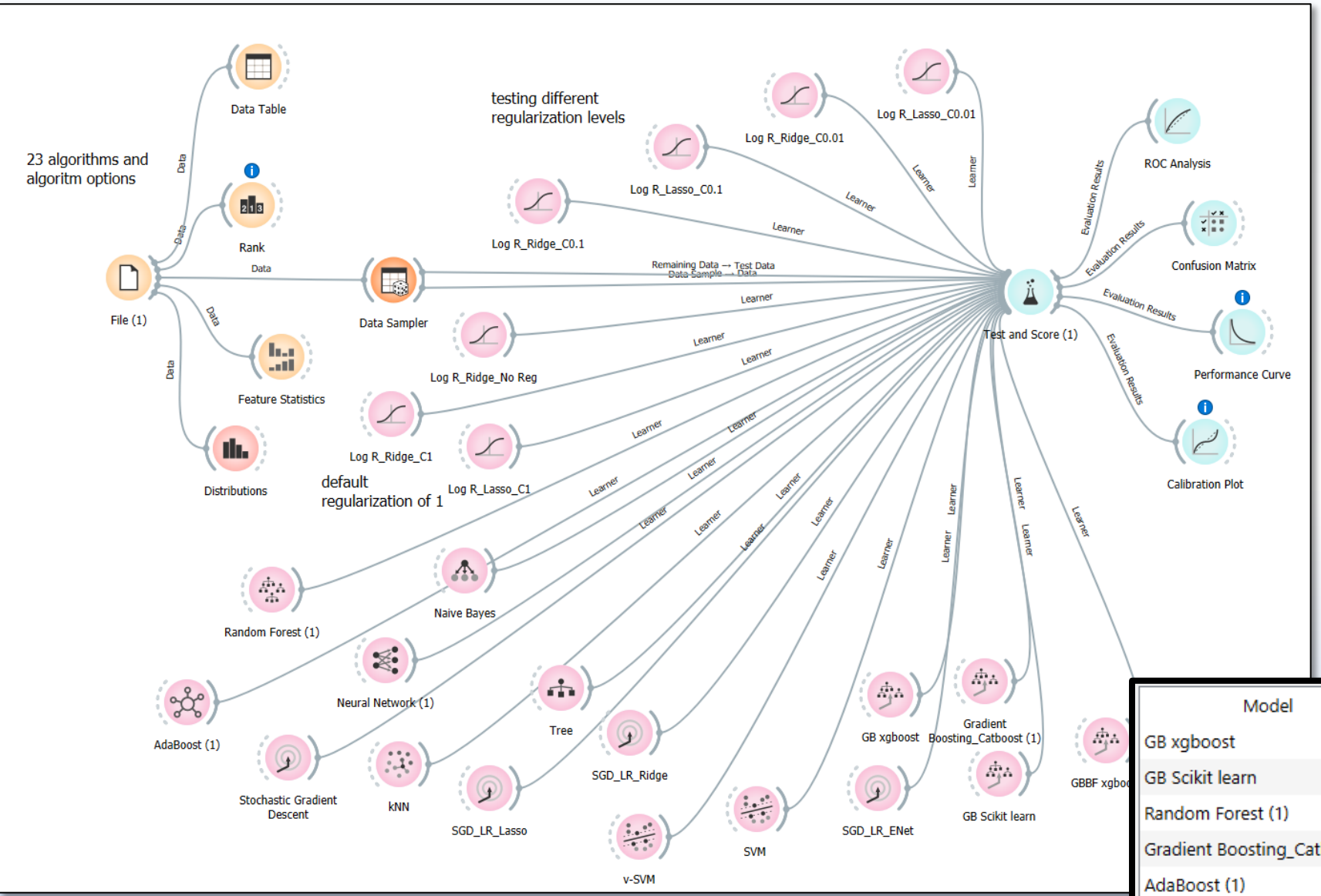
Simple predictive model:

This is a predictive model workflow for classification problems that compares 4 different algorithms or algorithm options

Test and Score

Model	AUC	CA	F1	Prec	Recall	MCC	Spec	LogLoss
Log R_Ridge_No Reg	0.857	0.788	0.803	0.828	0.780	0.575	0.798	0.458
Log R_Lasso_C1	0.857	0.788	0.803	0.828	0.780	0.574	0.797	0.458
Log R_Ridge_C1	0.857	0.788	0.803	0.828	0.780	0.574	0.797	0.458
Neural Network	0.856	0.760	0.785	0.780	0.791	0.513	0.721	0.522

Orange Visual Programming for Predictive Analytics and AI



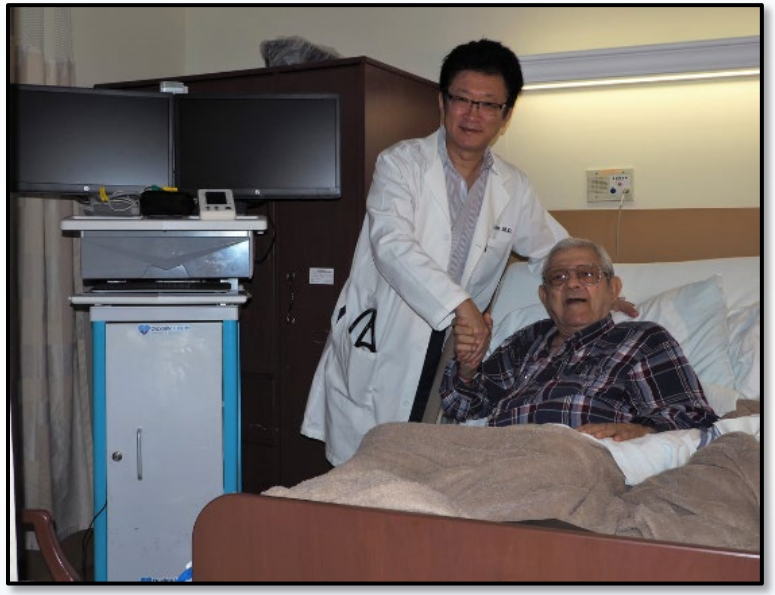
Complex predictive model: This is a predictive model workflow for classification problems that compares 23 different algorithms or algorithm options

Test and Score

Model	AUC	CA	F1	Prec	Recall	MCC	Spec	LogLoss
GB xgboost	0.998	0.990	0.992	0.990	0.993	0.981	0.987	0.036
GB Scikit learn	0.998	0.990	0.991	0.990	0.993	0.980	0.986	0.045
Random Forest (1)	0.999	0.990	0.991	0.989	0.993	0.979	0.986	0.051
Gradient Boosting_Catboost (1)	0.998	0.988	0.989	0.987	0.991	0.975	0.983	0.043
AdaBoost (1)	0.996	0.985	0.987	0.985	0.988	0.970	0.981	0.092
kNN	0.992	0.980	0.982	0.979	0.985	0.959	0.973	0.265
GBBF xgboost	0.983	0.968	0.972	0.972	0.971	0.935	0.964	0.464

Active Predictive Analytics Projects in our Pipeline

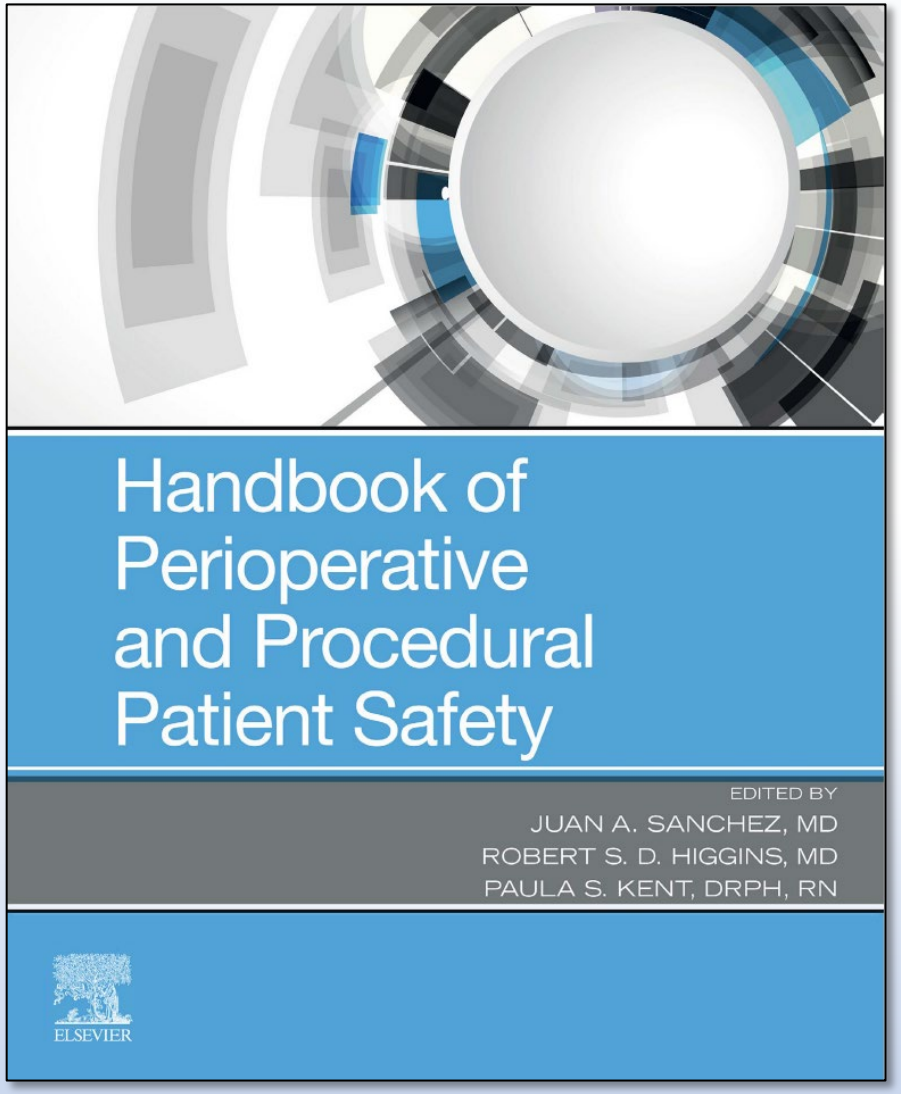
Michael Shen MD - Cardiologist
Founder and Chief Medical Officer, Duxlink Health
Miami-Fort Lauderdale Area



Jina Chung MD
Pulmonary Embolism Response Team
Division of Cardiology, Harbor UCLA Medical Center

H.E. (Skip) McCoy III
President-RiverSoft Home Healthcare Software Services
Melbourne, Florida

David Patrishkoff Co-authored a 45-page chapter in a 2023 John Hopkins-endorsed Handbook on Surgical Safety



CHAPTER

A perioperative safety and quality change management model and case study: Muda Health

15

Paul Barach, B.Med.Sci, MD, MPH, Maj (ret.), AUA[1,2], Hal Wiggin, PhD[3], Paul Risner, JD[8], Julie Johnson, MSPH, PhD[4], Dave Patrishkoff, MA[5], Shankar Kurra, MD, MBA[6], Becky Southern, RN[9], Edward Popovich, PhD[7]

[1]*College of Population Health, Thomas Jefferson University, Philadelphia, PA, United States;* [2]*Sigmund Frued University, Vienna, Austria;* [3]*Dr Koran C Patel College of Osteopathic Medicine, Health Informatics Department, Nova Southeastern University, Fort Lauderdale, FL, United States;* [4]*Department of Surgery, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States;* [5]*Dr Patel College of Osteopathic Medicine, Nova Southeastern University, Fort Lauderdale, FL, United States;* [6]*Sentara Virginia Beach General, Virginia Beach, VA, United States;* [7]*Dr Koran C Patel, College of Osteopathic Medicine, Nova Southeastern University, Fort Lauderdale, FL, United States;* [8]*President, Paul E. Risner, P.A;* [9]*Caldwell Butler Associates*

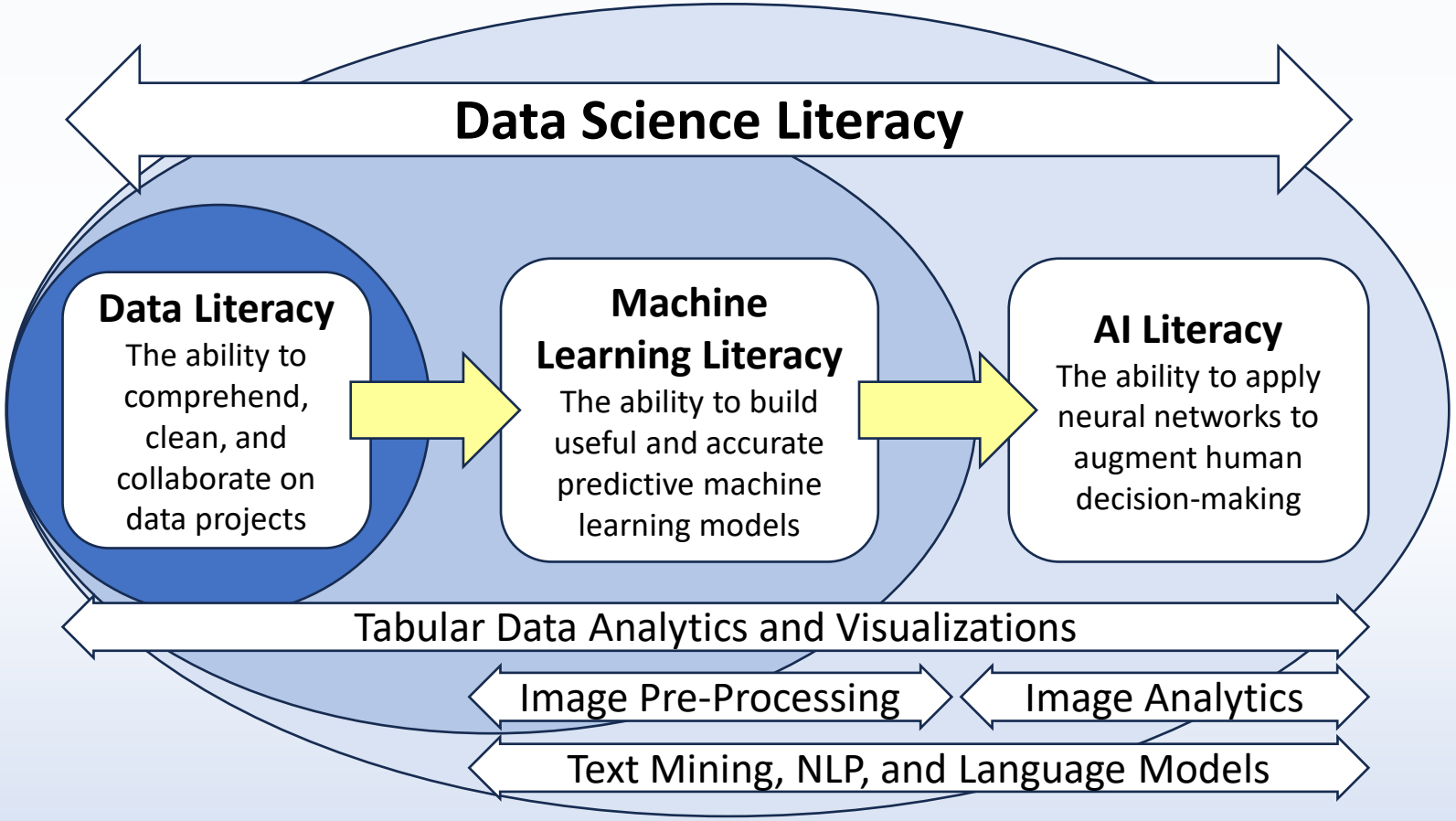
Every system is perfectly designed to get the results it gets.

— Paul Batalden, M.D.

Agenda

1. Introduction
2. What is Data Science and Predictive Analytics?
3. Potential Epic Applications of ML/AI
4. No code Predictive Analytics of Epic data

The Components of Data Science Literacy for Tabular Data, Image Analytics, and Text Mining



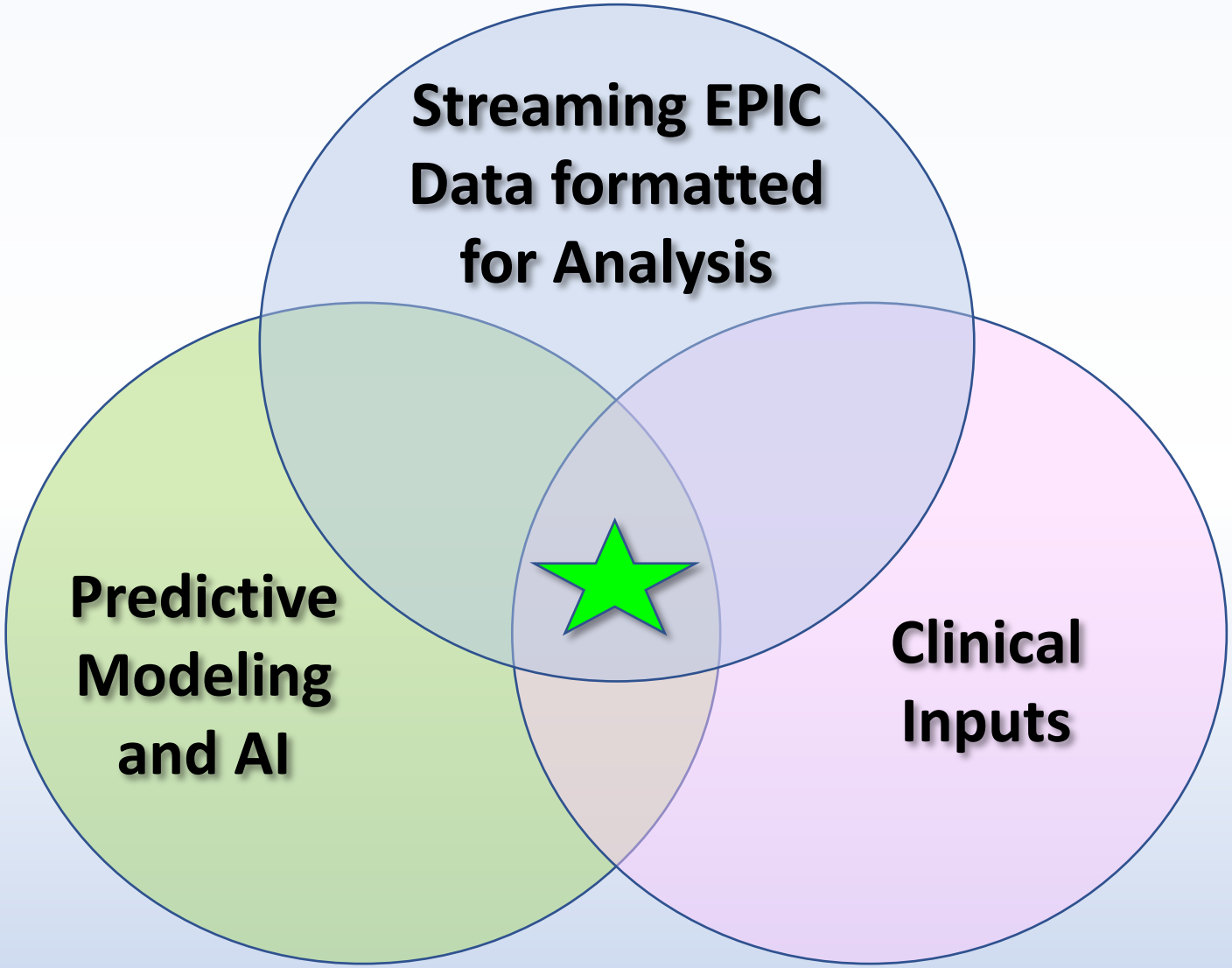
Data Science Literacy enabled with free open-source tools:
Orange, JASP, BlueSky Statistics, Google Teachable Machines, Lobe, and others

Comprehend, Interpret, and Communicate with Data >> Build accurate Predictive models >> Augment human decision-making

Agenda

1. Introduction
2. What is Data Science and Predictive Analytics?
3. Potential Epic Applications of ML/AI
4. No code Predictive Analytics of Epic data

3 Coordinated Activities to Analyze Epic Data



Using EHR Tabular Data for Predictive Analytics

1. Tabular Data (Structured Data)
    - **Patient Risk Scoring**: Use patient vitals, lab results, and past medical history to predict risks such as hospital readmissions, sepsis, or other complications.
    - **Patient Deterioration Predictions:** models can forecast if a patient is likely to deteriorate or need intensive care
    - **Chronic Disease Management**: For patients with chronic conditions like diabetes, trends in glucose levels or blood pressure can help forecast potential exacerbations
    - **Medication Management**: Predict which patients are likely to adhere to their medications or who may experience adverse drug reactions.
    - **Optimal Resource Allocation**: Predict patient flow to optimize bed management and resource allocation within the hospital.
    - **Cost Predictions**: Project future costs for patients based on their medical history, aiding in insurance and financial planning.
    - **And more**

Using EHR Images for Predictive Analytics

1. Images (Radiology, Pathology, etc.)
    - **Disease Detection**: Apply deep learning models to detect diseases in images like X-rays, MRIs, CT scans, and pathology slides. For instance, AI models can detect tumors in mammograms or lung nodules in chest X-rays
    - **Disease Progression**: Monitor the progression of diseases by comparing current images to historical ones. This is useful in conditions like multiple sclerosis or tumor growth
    - **Treatment Response**: Assess how diseases or conditions are responding to treatment over time by analyzing changes in images
    - **Dermatology**: Identification and classification of skin lesions, including malignant melanomas
    - **Ophthalmology**: Detection of diabetic retinopathy, glaucoma, or macular degeneration from retinal photographs
    - **And more**

Using EHR Text for Predictive Analytics

1. Text (Clinical Notes, Discharge Summaries)
    - **Natural Language Processing** (NLP): Extract relevant clinical information from physician or nurse notes, such as patient symptoms, treatments prescribed, or progress details.
    - **Disease Prediction**: Predict the onset of diseases based on textual symptoms described in clinical notes.
    - **Sentiment Analysis**: Understand patient feedback or concerns by analyzing text, helping hospitals improve patient satisfaction and experience.
    - **Clinical Decision Support**: Extract insights from the text to provide recommendations or alerts to clinicians, such as potential drug-drug interactions or overlooked symptoms.
    - **And more**

Agenda

1. Introduction
2. What is Data Science and Predictive Analytics?
3. Potential Epic Applications of ML/AI
4. No code Predictive Analytics of Epic data

The Need for More Data Scientists

Unleash the Data and its Potential with more NCDS Data Scientists and more Progress

Here are some of the reasons for the shortage of Data Scientists:

- Lack of high-end technical and non-technical skilled Data Scientists in the market
- Data Science skills have a steep learning curve with coding and programming being one of the biggest roadblocks for entry in this field
- Data Science programs are often only offered to graduate programs which limits entry for a broader pool of talent into this industry
- Many large companies compete with each other to implement ML/AI which consumes most of the available Data Scientists

Dilemma:

- Most small and mid-sized organizations cannot afford Data Scientists to help them with ML/AI implementation

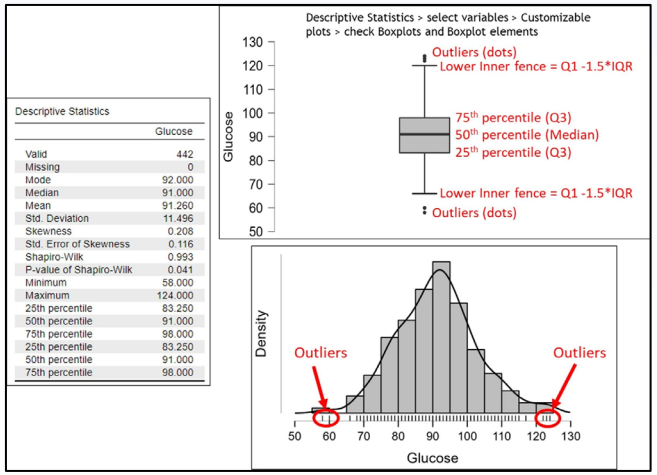
Adding new trained and certified **No-Code Data Scientists** to the market could help to close the skills gap and avoid the coding roadblock in the learning process

No-Code Data Science Capabilities

1. Conducting Data preparation and wrangling
2. Constructing compelling data visualizations
3. Developing and evaluating predictive models
4. Conducting Time Series forecasts and survival analysis
5. Conducting Geolocation-based analysis
6. Implementing Image Analytics Techniques
7. Mastering Text Mining Strategies
8. Exploring the future of Continuous Improvement Methodologies

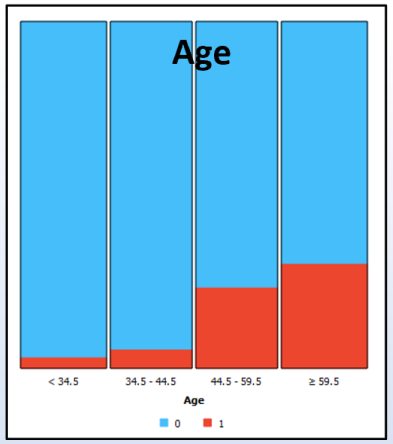
1. Data Preparation and Wrangling

Reshaping Chaos into Clarity: Polishing and Prepping Data for Predictive Modeling

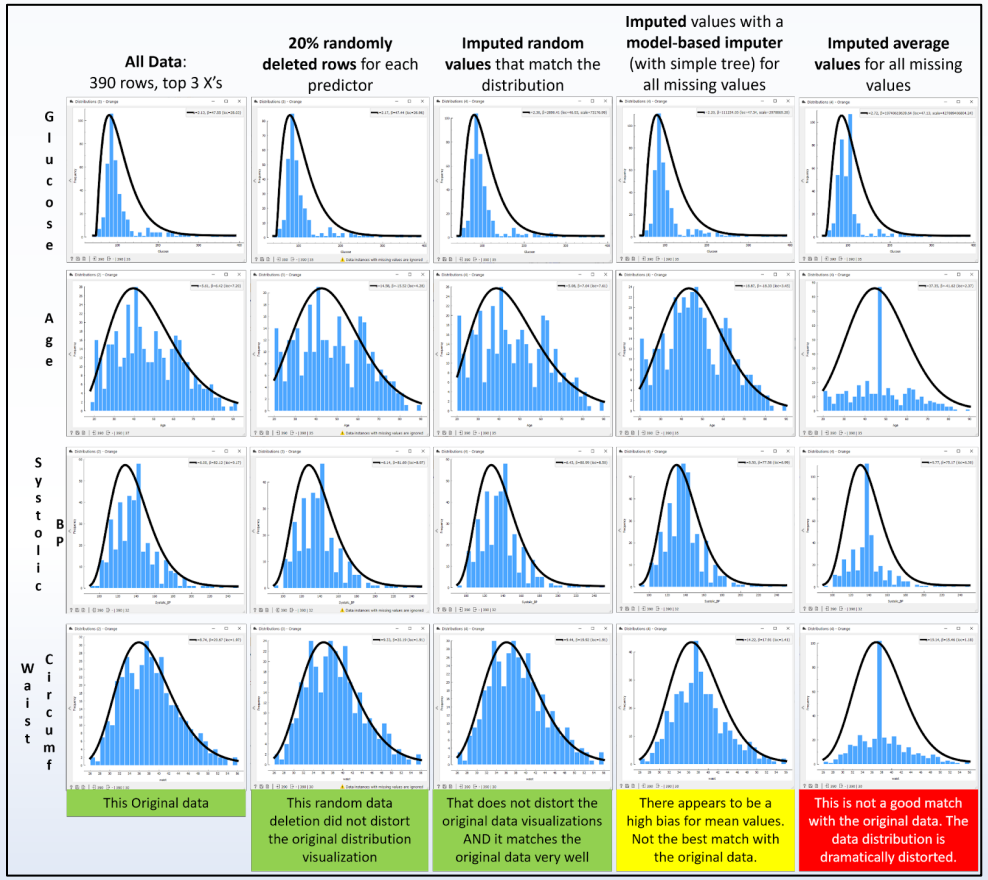


Descriptive Statistics	Glucose
Valid	442
Missing	0
Mode	92.000
Median	91.000
Mean	91.260
Std. Deviation	11.496
Skewness	0.208
Std. Error of Skewness	0.116
Shapiro-Wilk	0.993
P-value of Shapiro-Wilk	0.041
Minimum	58.000
Maximum	124.000
25th percentile	83.250
50th percentile	91.000
75th percentile	98.000
25th percentile	83.250
50th percentile	91.000
75th percentile	98.000

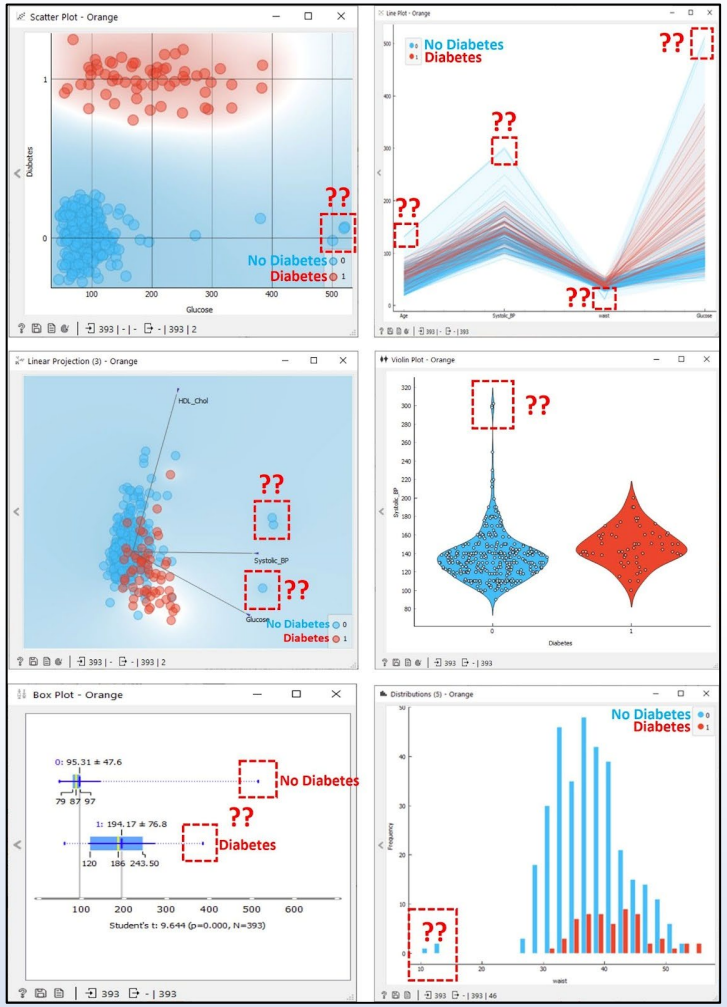
JASP Outlier analysis



Orange data binning



Orange Missing Data Imputation experiments

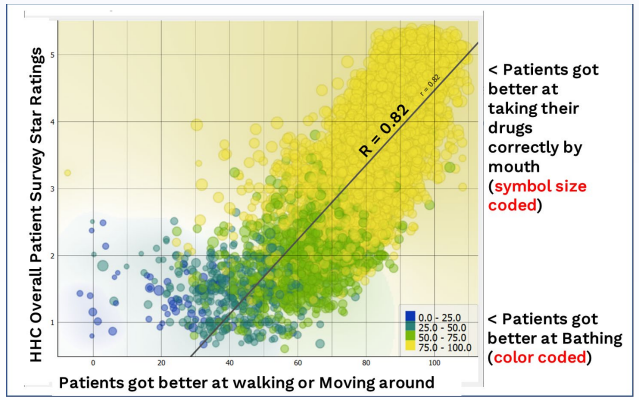


Orange data entry error identification

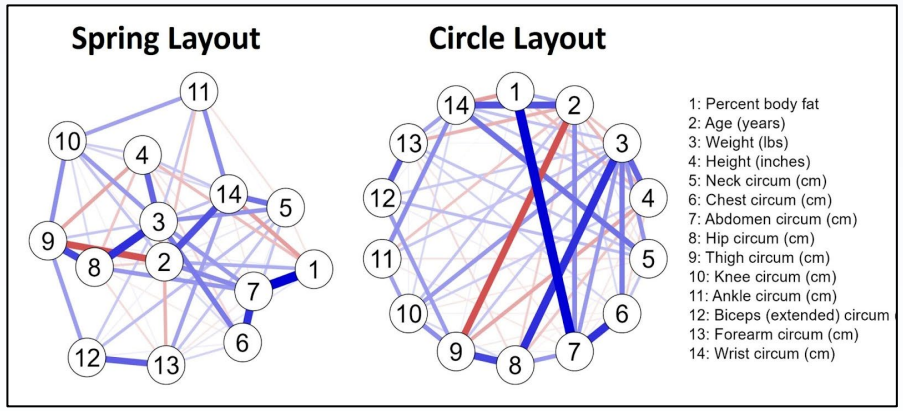
Orange, BlueSky, and JASP offers a wide variety of data prep and wrangling techniques

2. Constructing Compelling Data Visualizations

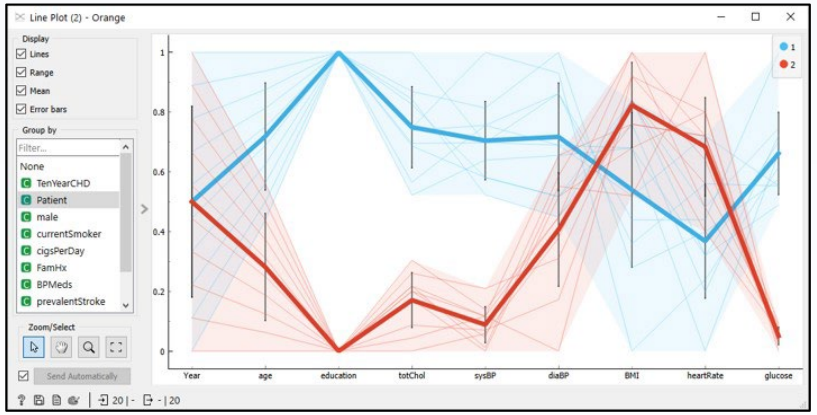
Painting Stories with Numbers



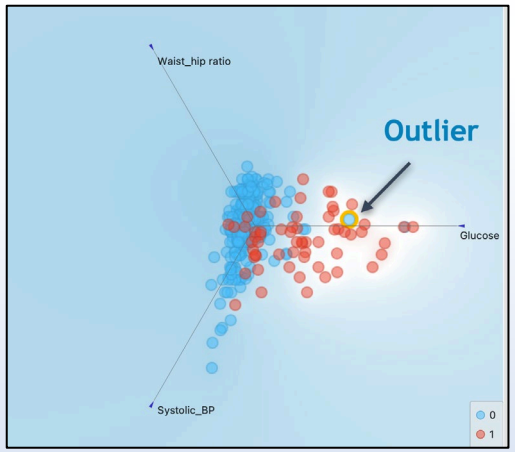
Orange Scatter Plot



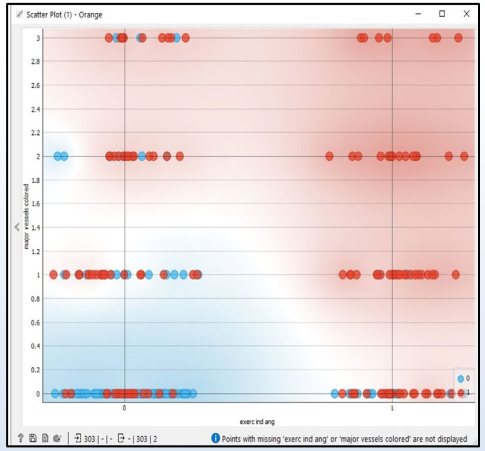
JASP Network / Correlation plots



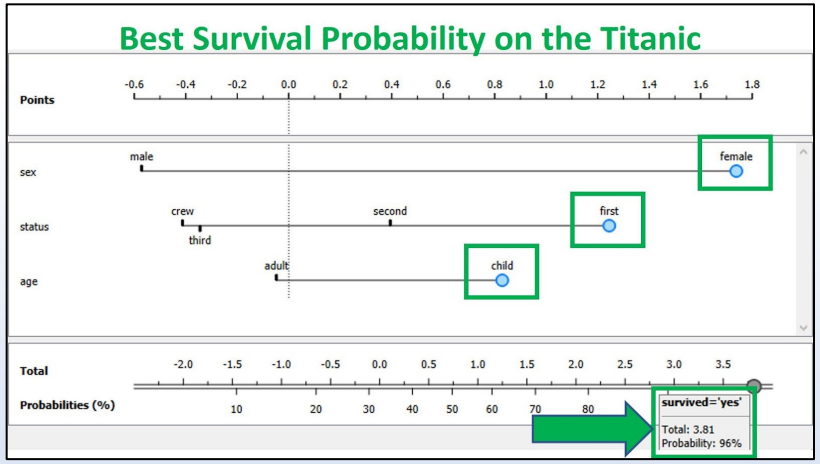
Orange Parallel Coordinates Plot



Orange 3D plot



Orange Scatter Plot

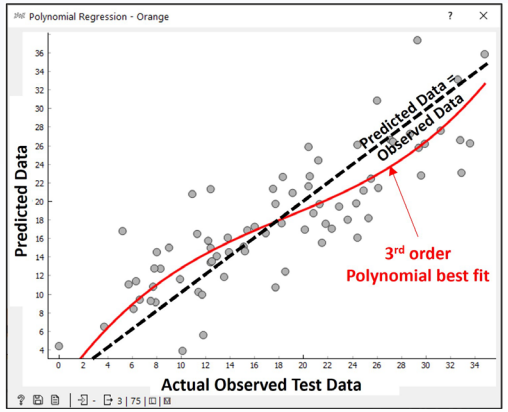


Orange Nomogram

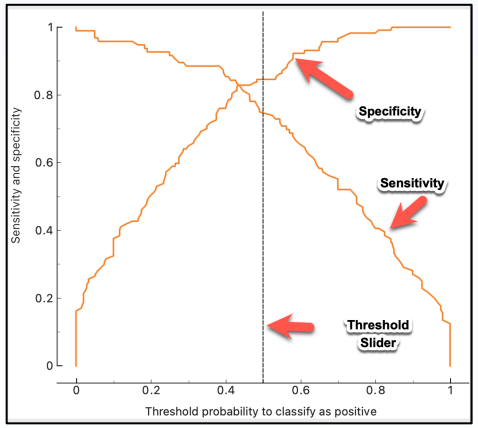
Orange, JASP, and BlueSky Statistics offers a wide variety of data visualizations

3. Developing and Evaluating Predictive Models

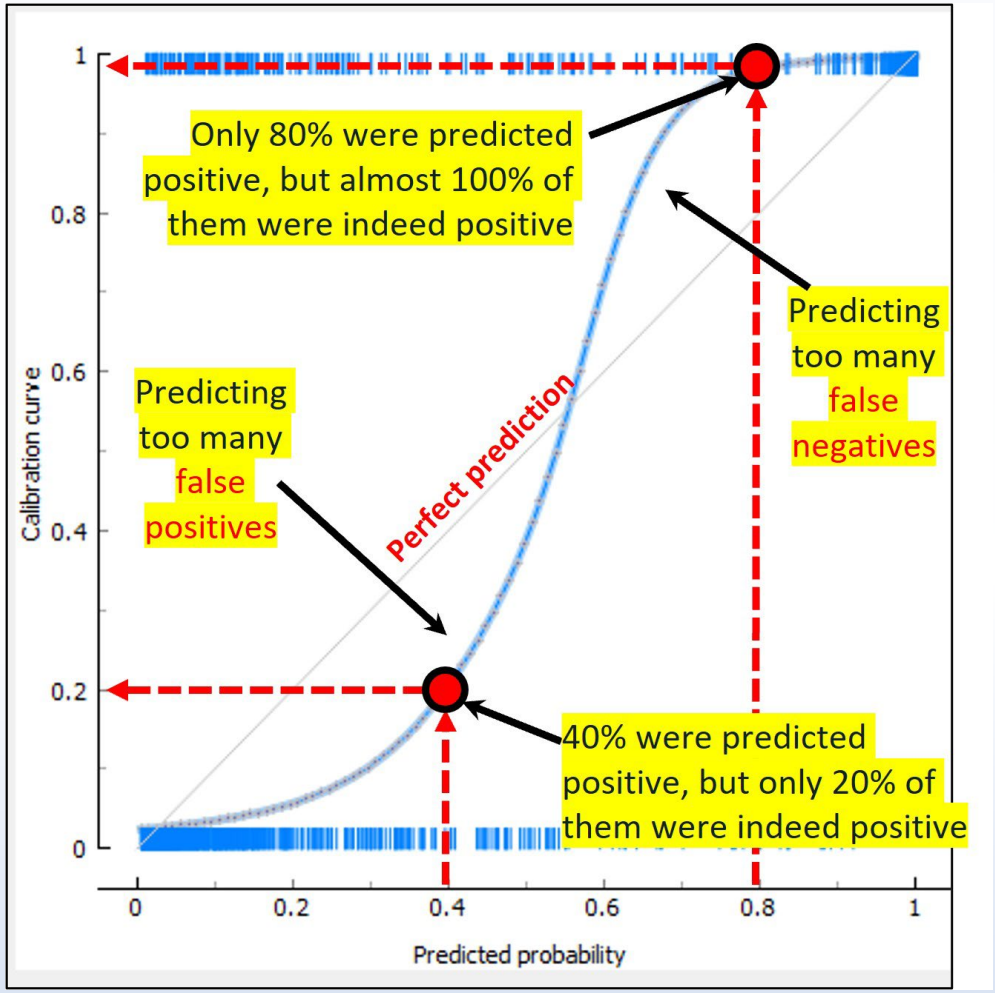
Forging Future Insights: Crafting and Calibrating Predictive Models



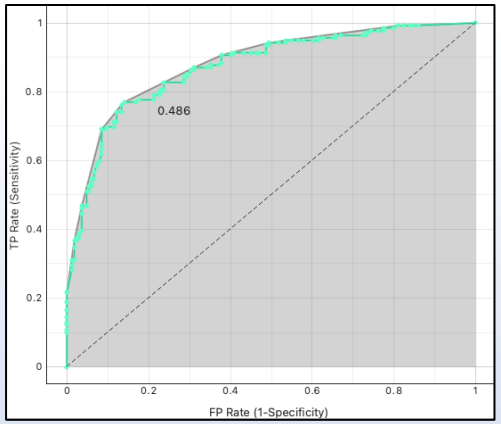
Orange Model Residuals analysis



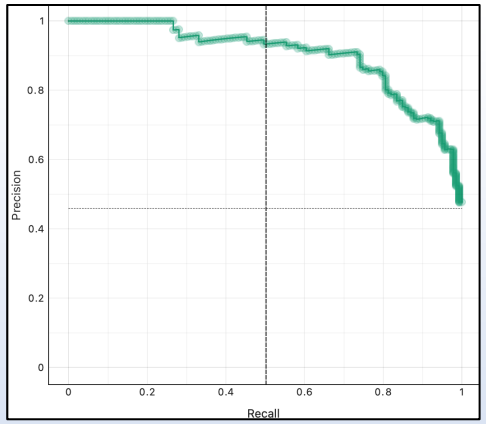
Orange Sensitivity vs Specificity plot with Threshold slider



Orange model Calibration Curve



Orange ROC Curve

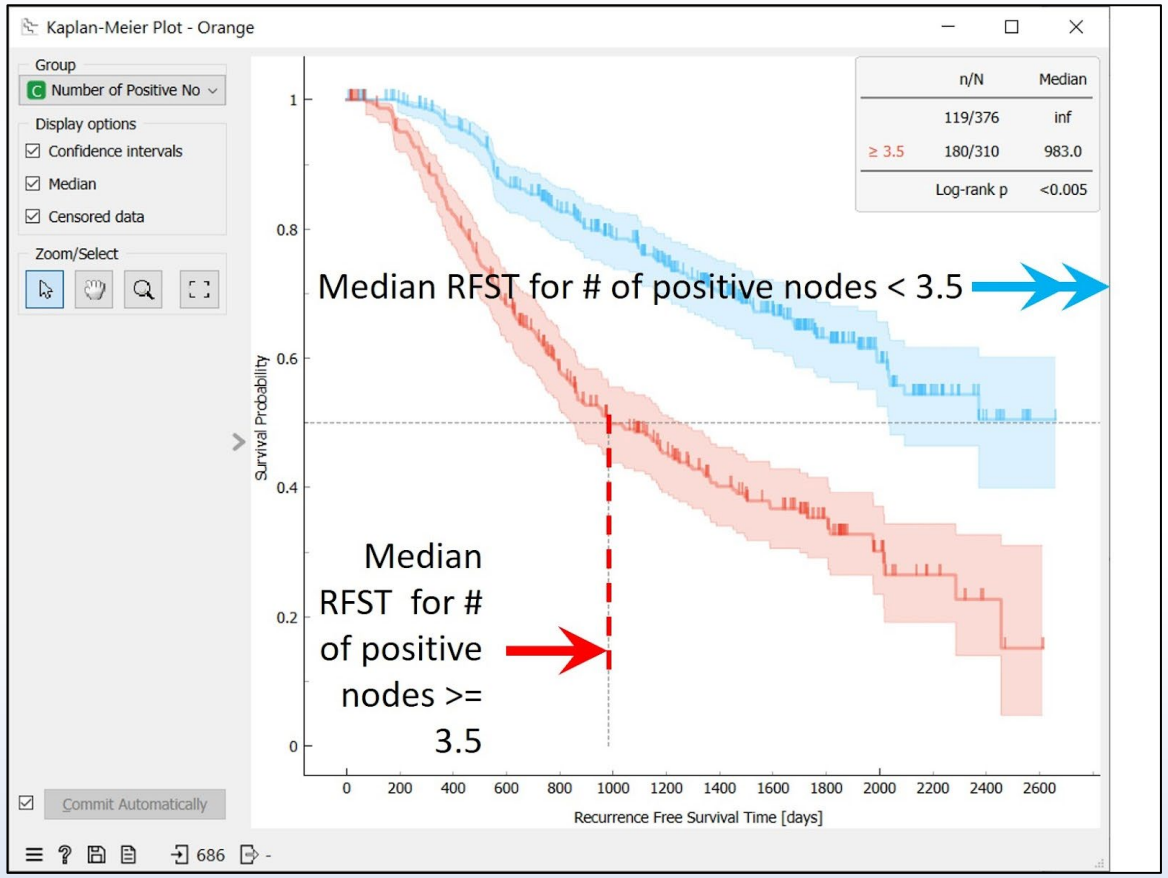


Orange Precision-recall Curve

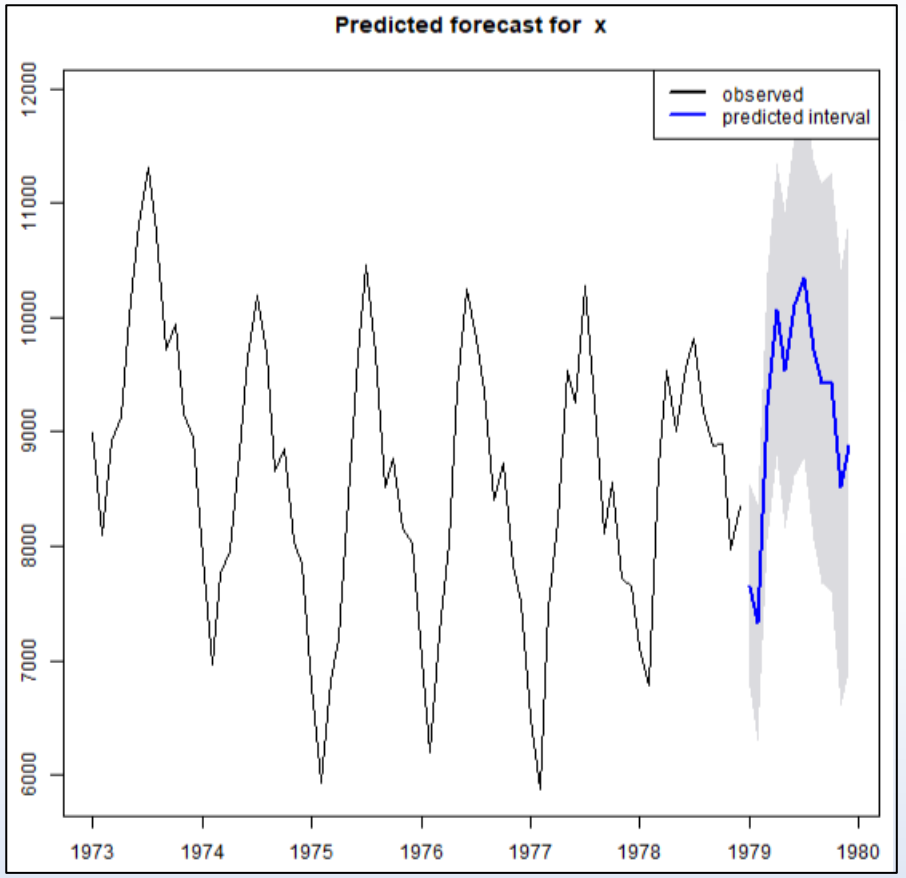
Orange, BlueSky, and JASP offers a wide range of model building and evaluation methods

4. Conducting Time Series Forecast and Survival Analysis

Unraveling Time's Tapestry: Mastering Forecasts and Lifelines with Analysis



Recurrence times for breast cancer

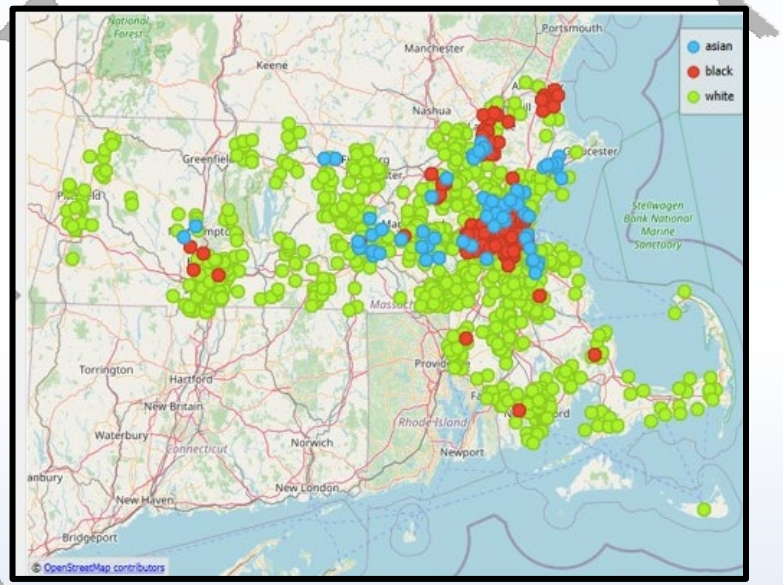


Orange and BlueSky offers Time Series Forecasting and Survival Analysis techniques

5. Conducting Geolocation-based Analysis

Merging Art, Geography, and Analytics

Analytical



Geospatial Visualizations

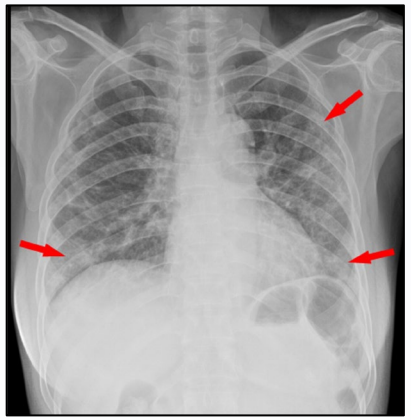
Creative

AI Images created by DALL-E

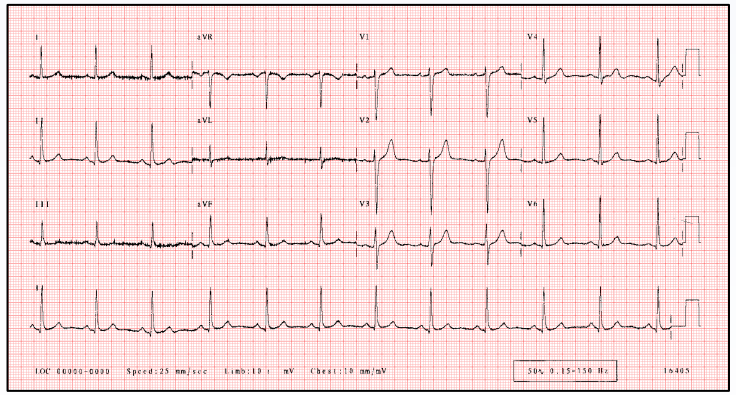
Orange and BlueSky offers Geo-Spatial Analysis techniques

6. Implementing Image Analytics Techniques

Picturing the Future: Predictive Insights with Image Analytics

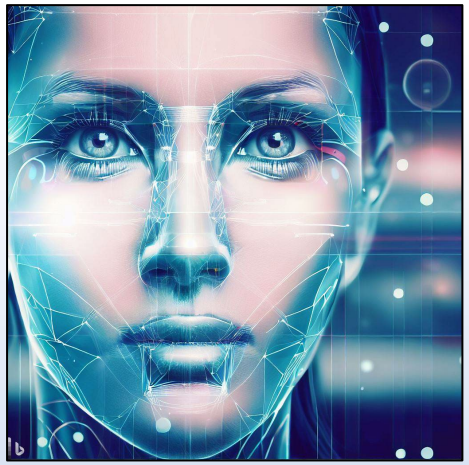


Xray diagnoses

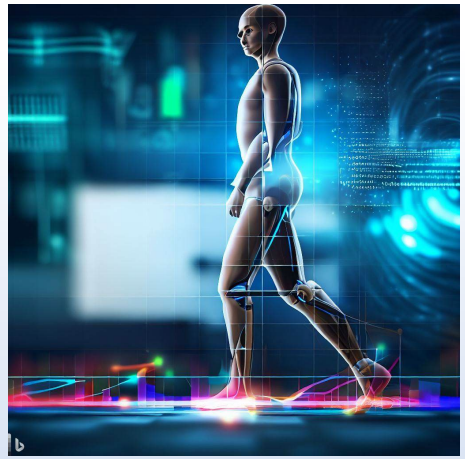


ECG diagnosis

Mood detection



Facial recognition



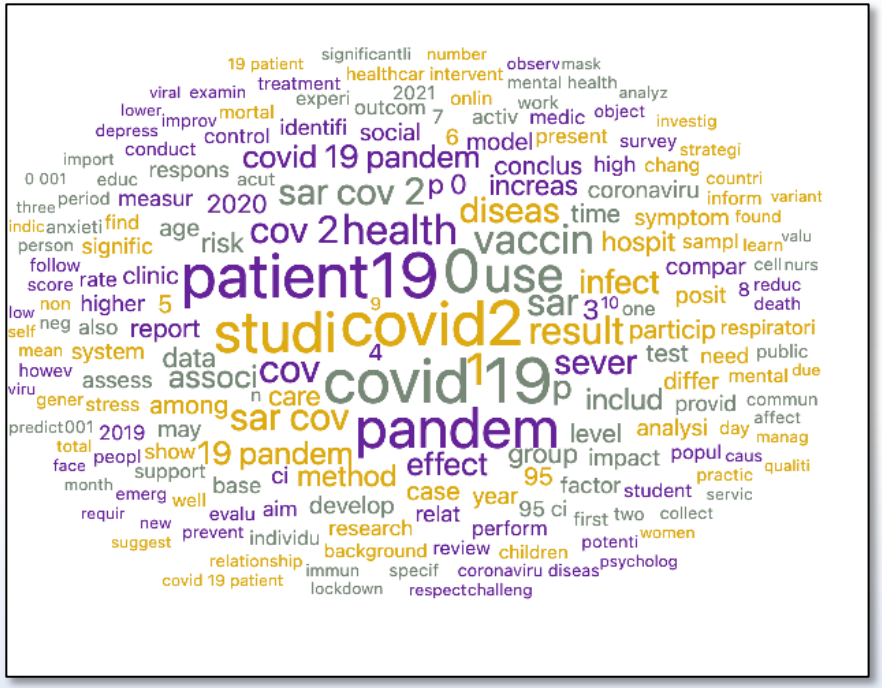
Posture / gait classification

AI Images created by DALL-E

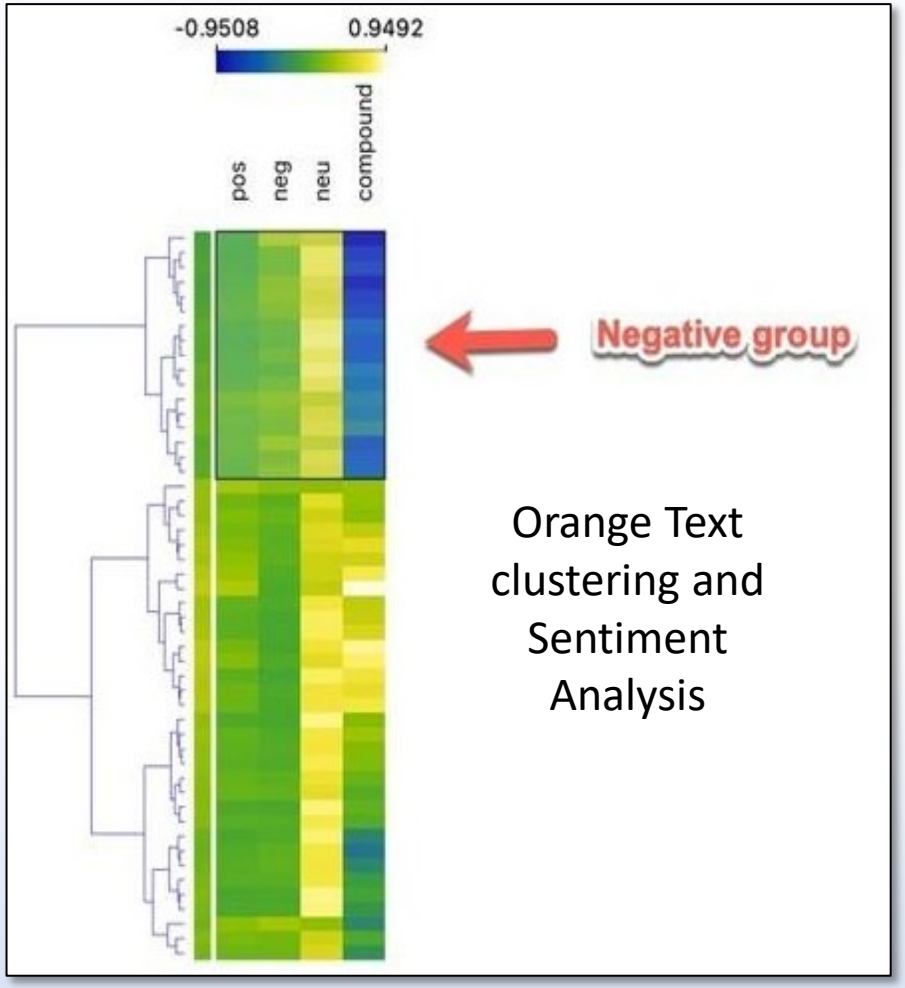
Orange can provide a wide range of image detection, classification, and predictive analytics

7. Mastering Text Mining Strategies

From Words to Wisdom: Master the Art of Text Mining for Deep Insights



Orange Word Cloud

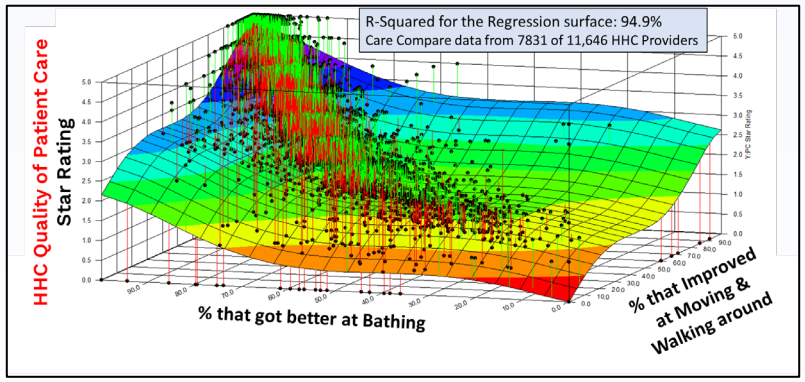


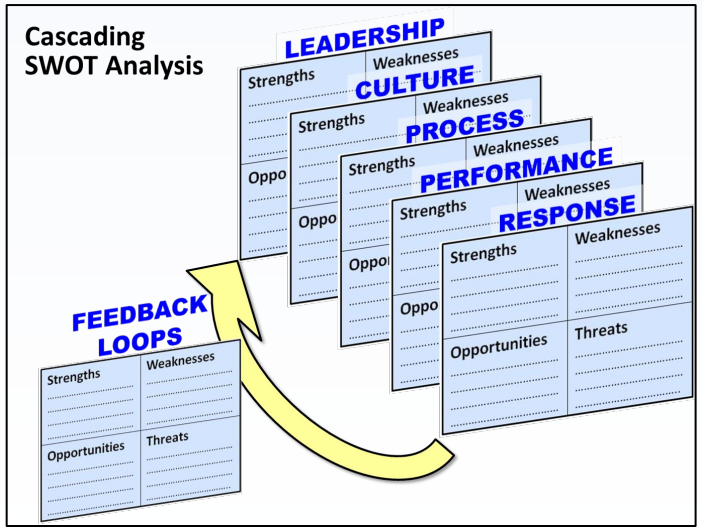
Orange Text clustering and Sentiment Analysis

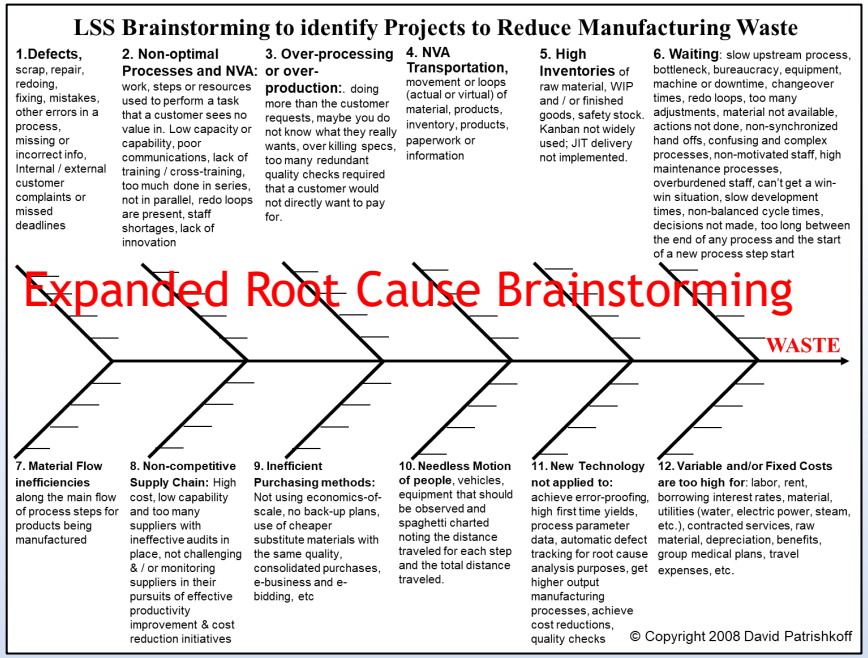
Orange offers a variety of Text mining tools ranging from simple Word Clouds to Text Clustering, Sentiment Analysis, and Text Grouping Predictions

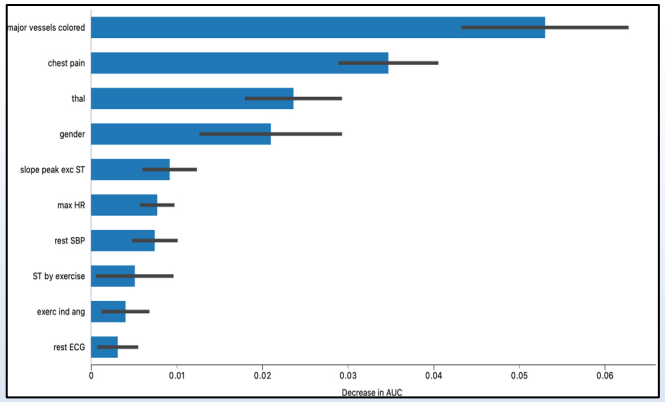
8. Exploring the Future of Continuous Improvement (CI) Methodologies

Revolutionizing Efficiency: Disruptively Reinventing Lean Six Sigma and other CI Methodologies with ML/AI









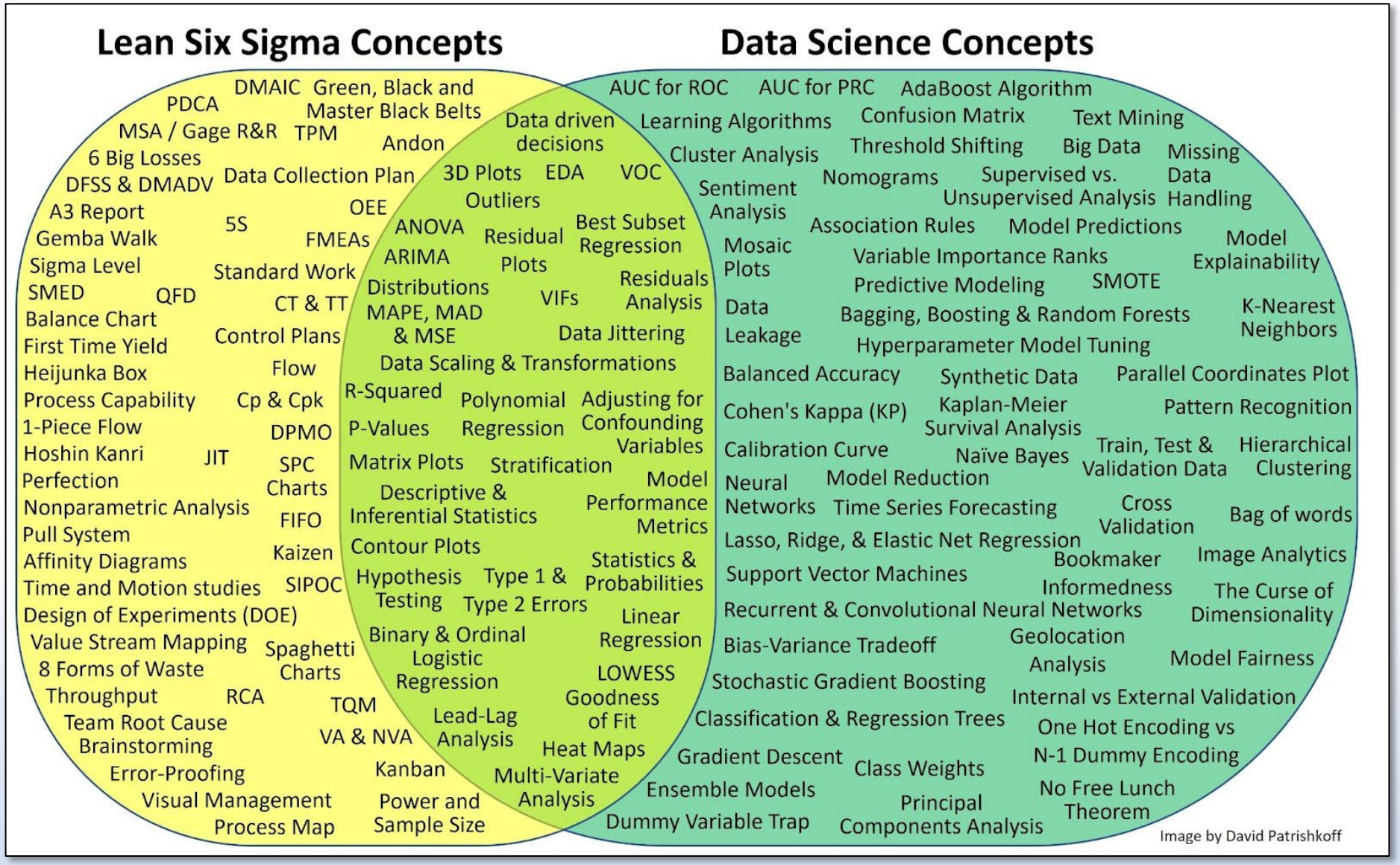
Orange Feature importance results for a classification problem

The future of CI will focus on:

- Realtime and automated data analysis at the Gemba
- Expanded cascading root cause analysis
- Rapid continuous improvement
- Organizational cultural risk identification and mitigation
- AI chatbot support for all problem brainstorming activities
- Integrated Predictive Analytics
- ChatPDF access to improved processes and procedures

BlueSky Statistics covers every standard and advanced Six Sigma Data Analysis Technique

Lean Six Sigma (LSS) and Data Science Vocabulary



LSS professionals are strategically positioned to be great candidates for No Code Data Science training

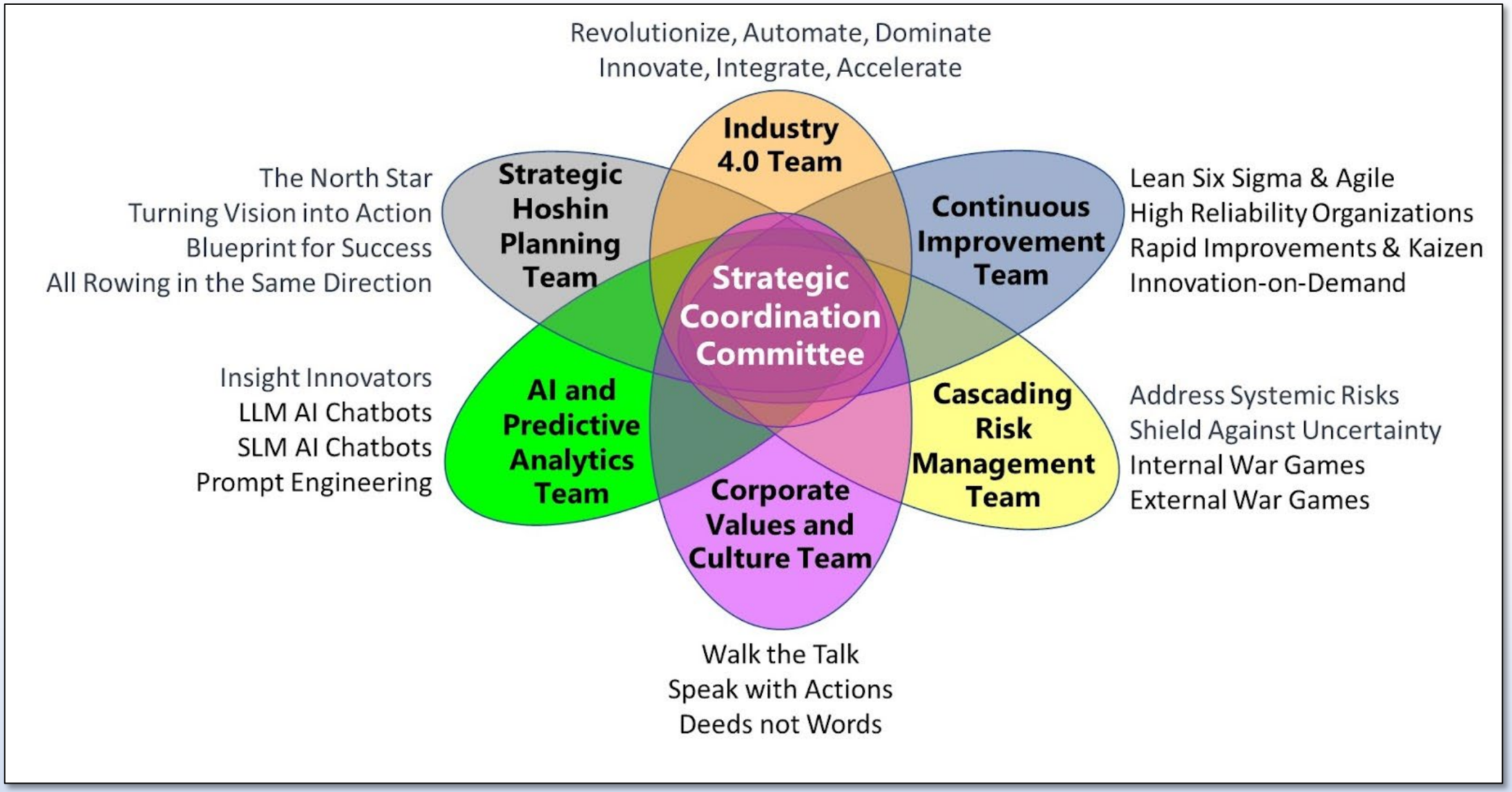
AI Chatbots can Dramatically Augment the Performance of Clinicians, Professionals and Continuous Improvement Projects

- AI Chatbots provide instant and valuable inputs for personal and CI teams, especially when they have been properly provided with background and goal "**prompts**"
- **Prompt engineering** is a new field and profession that is the art and science of designing prompts (requests) for AI chatbots in order to maximize the value of the AI chatbot outputs
- "**Super prompts**" have been shown to create outputs that are equal to the best Innovation-on-demand (TRIZ) experts
- Our own testing has shown superior outputs for the following areas when situational awareness is provided: **SWOT analysis**, **FMEA analysis**, **Cause and Effect Fishbone brainstorming**, **Solution matrix brainstorming**, **Project Data collection brainstorming**, **process maps in text format**, and more

AI Chatbots can Dramatically Augment the Performance of Clinicians, Professionals and Continuous Improvement Projects

- We are also experimenting the effectiveness of AI chatbots to **instantly** create the following reports from **doctor** and **nursing notes** that are available in EHRs:
  - **SOAP** (Subjective, Objective, Assessment, Plan)
  - **SOAPIE** (Subjective, Objective, Assessment, Plan, Intervention, Evaluation)
  - **DAR** (Data, Action, Response)
  - **SBAR** (Situation, Background, Assessment, Recommendation)
  - **ISBAR** (Identify, Situation, Background, Assessment, Recommendation)
  - And other clinical reports (out-patient notes, in-patient notes, etc)

ML and AI are at the Core of Industry and Healthcare 4.0



The Redwood Approach: Building Business Resilience Through Cooperative Strength

- Redwoods stand over 350 feet and live for 2,000 years, drawing strength from interconnected roots that intertwine up to 100 feet away
- Fairy circles arise as young trees sprout from a parent's roots, sharing resources
- Imagine a business as a redwood cluster: each tree symbolizes a segment of a strategy for Healthcare 4.0
- Interwoven strategies boost resilience and growth against industry challenges
- ML/AI-driven predictive analytics can be based on Data sharing between departments through Epic

Conclusion

- New potential Insights in **Epic** data can be unlocked with no code predictive analytics tools and free open-source software
- ML and AI will create the new and expanded opportunities to improve healthcare outcomes when access to **Epic** data is maximized
- AI chatbots can create summarized doctor and nursing notes in **Epic** for each patient in any desired standard format
- Continuous improvement programs will be disrupted by AI Chatbots with their ability to provide valuable and instant brainstorming inputs, if correctly prompted

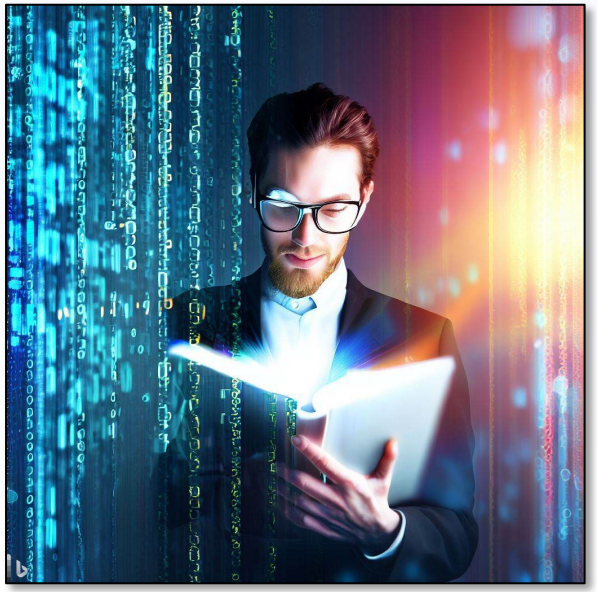
Attachments – Additional Information

Biographical Sketches

David Patrishkoff, M.S. is a Lean Six Sigma Master Black Belt with C-level worldwide executive experiences with engineering, quality, and manufacturing responsibilities at multi-billion-dollar revenue companies. In 2001, he founded a consulting and training company, E3 (www.e3.business), where he has trained and consulted for organizations in over 60 different industries worldwide to resolve their mission-critical issues with innovative process improvement and data analysis techniques. In 2018, he added machine learning techniques to gain more insights into his research of large datasets, which included the Fatality Reporting Analysis System (FARS) for motor vehicle accidents in the USA and the Home Health Outcome Assessment Information Set (OASIS) data for home healthcare patients in the USA. He speaks regularly at international conferences about his research and other topics of interest. He is a contributing faculty member at the Kettering University School of Management in Flint, Michigan. He is also an Adjunct Professor at the Dr. Kiran C. Patel Osteopathic School of Medicine in Ft. Lauderdale, Florida, part of Nova SE University and an Adjunct Professor at SUNY Polytechnic Institute in Utica, NY. He has already incorporated Orange into some of the analytics classes he teaches at Nova University and the SUNY Polytechnic Institute. He has trained, certified, and mentored many professionals in various topics, including over 3,000 professionals in Lean Six Sigma techniques. He and his company have also trained over 23,000 healthcare professionals in High-Reliability Organizations (HRO) plans to reduce medical errors in healthcare systems.

Dr. Robert Hoyt FACP FAMIA ABPM-CI is an internal medicine physician with a well-rounded background. He was in private practice for fifteen years in Virginia and then returned to active duty, completing three years in the US Army and seventeen in the US Navy, and retiring at the rank of Captain. In the last several years on active duty, he conducted research at the Navy Aerospace Medical Research Laboratory (NAMRL) and the Robert E. Mitchell Center for Prisoner of War Studies, located in Pensacola, Florida. He created and taught health informatics at the University of West Florida for thirteen years. He is the editor and author of Health Informatics: Practical Guide, which is in its eighth edition. In 2014, he became board-certified in clinical informatics, and in 2016 he became a Fellow of the American Medical Informatics Association (FAMIA).
Since 2018, he has focused his attention primarily on data science as a new field that has impacted every domain, including medicine. He is passionate about teaching clinicians about important new trends in informatics and data science. This led to the creation of two new textbooks, Introduction to Biomedical Data Science (2019) and Data Preparation and Exploration (2020). Information about those textbooks can be found at https://www.informaticseducation.org . Dr. Hoyt is the honorary president of the Medical Intelligence Society, contributing monthly to the Data Science Tip of the Month virtual meetings. He is on the Board and faculty for the American Board of Artificial Intelligence in Medicine Board Review Course https://abaim.org/ . Furthermore, he is a reviewer for multiple medical journals, including the new Intelligence Based Medicine journal https://www.sciencedirect.com/journal/intelligence-based-medicine .

What is Data Literacy?



Delve into data basics to understand its essence and role

Acquaint yourself with data formats and their organization

Tackle visualization by interpreting and creating data visuals

Apply statistical fundamentals for insightful data analysis

Learn about data cleaning, refining it for better analysis

Internalize ethics, recognizing our moral duties in data handling

Train with tools, mastering the essentials of data software

Express insights through storytelling, crafting compelling narratives

Research deeper into data science and its advanced techniques

Adventure through projects, applying your skills in real-world scenarios

Cultivate a systematic approach with a clear discovery process

Yield mastery and confidence, attaining proficiency in all data challenges

What is Data Science (DS) Literacy?

DS Literacy based on our DISCOVER Process:

Define Research questions, Problem Statement, and Goals for the Project

Integrate, Collect and import High Quantity and High-Quality Data

Scrub, clean, transform, feature engineering, dummy encoding, and check multi-collinearity (VIFs)

Cultivate preliminary Insights with Exploratory Data Analysis

Optimize model performance with experiments and model tuning

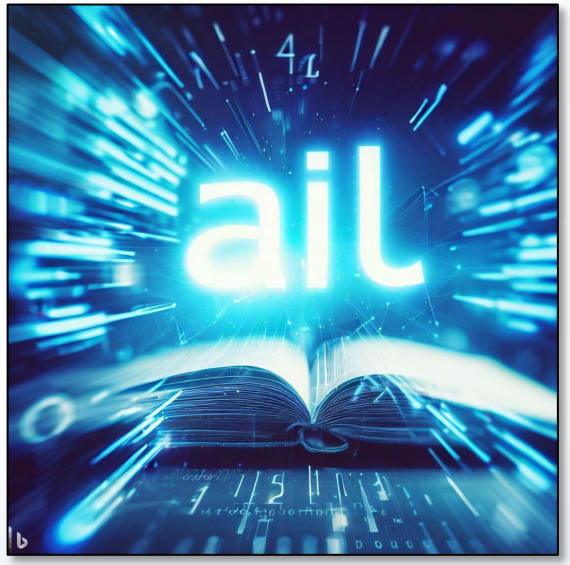
Verify model performance with various model evaluation methods

Explain model insights effectively to stakeholders to build trust

Reevaluate and Retrain the model when model drift or concept drift is confirmed

What is AI Literacy

The AI GOAL as a structured path to AI Literacy:



A.I essentials: delve into the core concepts, definitions, and history of Artificial Intelligence

Grasp algorithms and the diverse techniques that are powering AI solutions

Operate with Tools and various AI platforms for hands-on learning

Amass knowledge in Application Awareness to see AI's real-world impact

Learn about fairness, ethical, bias-free, moral, societal, and regulatory considerations of AI