

Case study of verification and validation methods for algorithmic predictive biomarkers

Kartik Shetty, Priya Gopal
Kartik.shetty@takeda.com, priya.gopal@takeda.com
Clinical Data Engineering



- **Biomarkers:**
 - Measures taken at a specific time on cells or organisms
 - Indicates disease, infection, or environmental exposure
 - Focuses on predictive biomarkers for forecasting health conditions
- **Efficient Data Gathering: (Wearable devices, Predictive biomarkers)**
 - Wearable devices efficiently collect extensive healthcare data
 - Enables the utilization of software and technologies to generate predictive biomarkers
- **Digital Health Technology Impact:**
 - Evolution of wearable devices
 - Prioritizing patient-centricity for enhanced ease of use and accessibility
- **Application of predictive biomarkers**
 - Used in a clinical trial setting to accelerate drug development and patient care
 - Must meet clinical research standards for such application

Case study: Welltory COVID-19 Dataset



- **Objective: Create a Predictive Biomarker**

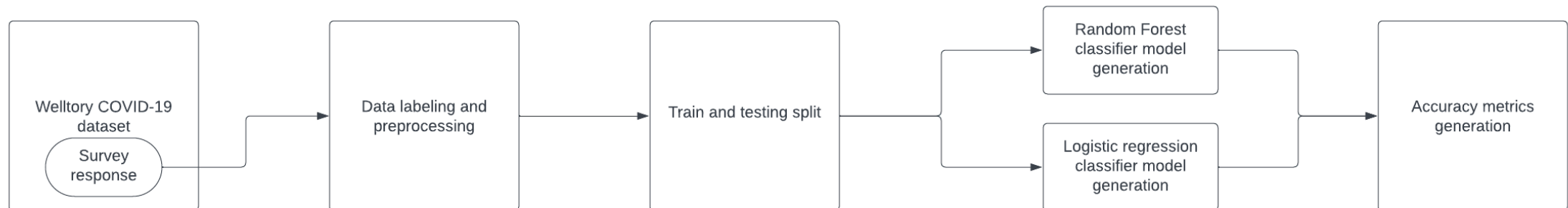
- Anticipate severe COVID symptoms in positive cases using wearable device data
- Identify different steps involved in the process of creation
- Explore verification and validations methods for the steps identified

- **Dataset**

- Welltory Dataset curated for COVID-19 research in 2020
- Information included heart rate variability, wearable device data and clinically verified assessments of physical and mental health
- Dataset comprised data from 68 patients, with a size of 24 megabytes

- **Predictive biomarker generation**

- Utilized Random Forest and Logistic Regression classification techniques
- Due to limitation in size of the dataset, generated synthetic data to train



Accuracy metrics



- **Accuracy:** ability of model to distinguish between the two classes, with higher scores indicating better overall performance
- **Precision:** the proportion of correctly identified positive instances among all instances predicted as positive
- **Recall:** the proportion of correctly identified positive instances among all actual positive instances
- **F1-score:** combines both precision and recall into a single value
 - balance trade-off between precision and recall

RANDOM FOREST ALGORITHM

LOGISTIC REGRESSION ALGORITHM



Accuracy	Precision	Recall	F1-score
0.85	1.0	0.6	0.75

Accuracy	Precision	Recall	F1-core
0.75	0.75	0.45	0.56

The improvement in performance is evident when considering all the metrics for both the models.

Conclusion: It is important to have an **even distribution** of both the positive and negative labeled datapoints.

Verification & Validation of Algorithmic Biomarkers



- Demonstrate safety, effectiveness, and performance of the biomarker
- Verification and validation must be conducted to for use in a real-world setting
- **Verification:** software meets the specified requirements and design specifications
 - provides support to prove that software is validated
- **Validation:** software meets the user's needs and is fit for its intended purpose
- Thorough clinical evaluation must be carried out: **Valid clinical association, Analytical validation, Clinical validation**



Valid clinical association

Is there a valid clinical association between output and target condition? By using existing evidence(Literatures, research etc.) or generating new ones(data analysis)



Analytical validation

Is the software built the right way, meet its requirements and intended use?



Clinical validation

Does the output data achieve the intended purpose in the context of health care?

Analytical Validation



- Confirm the accuracy, reliability, and precision of the software
- Provide evidence that software is performing as intended
- Ensure software produces accurate and reliable results

RECOMMENDATIONS

- 1 Define scope, requirements and limitations of the software. In the context of the case study, the predictive biomarker predicts the severity but does not predict the symptoms itself
- 2 Perform validation on each of the steps identified in case study separately using high level guidance bullet points provided by regulatory agency for different activities involved in the software development lifecycle(SDLC)
- 3 Data collection is validated for multiple independent cohorts to ensure generalizability and performance of the algorithm against different patient populations
- 4 Utilize accuracy metrics for validating the effectiveness of the algorithmic biomarker in predicting precise outcomes. In the case study, we employed metrics such as F1 score, precision, and recall to assess and confirm the algorithm's performance.
- 5 Set threshold for the metrics to ensure and benchmark performance

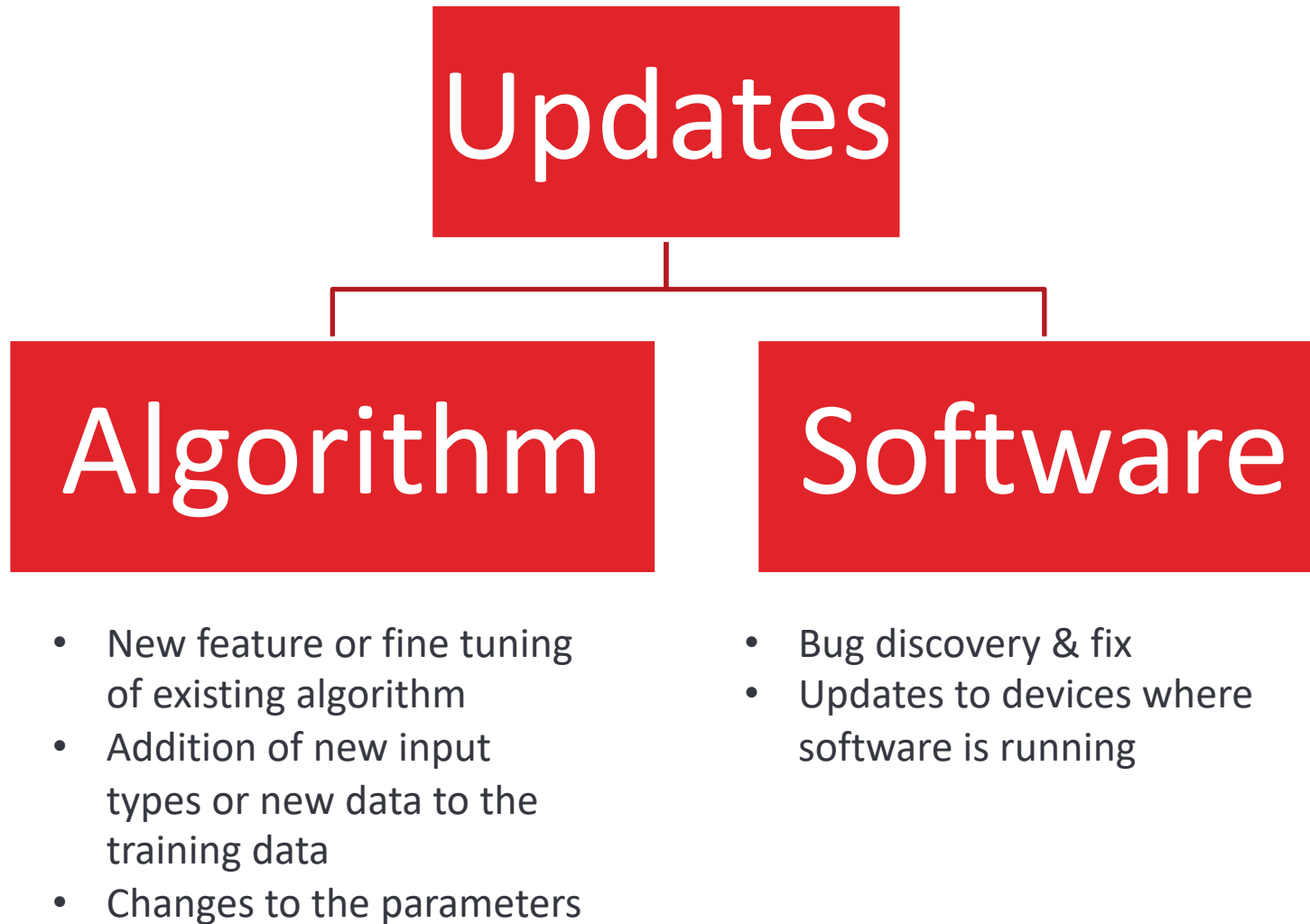
Clinical Validation



- Assessing biomarker's clinical relevance and predictive value
- Testing the biomarker in well-designed clinical studies

RECOMMENDATIONS

- 1 Compare prediction by the biomarker against an alternative method. In the context of the case study on heart rate variability measurement, compare the traditional measurement of heart rate with the output measured by the device.
- 2 Test algorithm across diverse set of parameters - diverse demographic patient population and sex), with/without pre-existing conditions, ensuring reliability and effectiveness
- 3 Prediction certified by the clinician to demonstrate clinical accuracy. In the context of the case study, they can compare the output of the biomarker with their own diagnosis to assess accuracy



Handling updates and changes: Analytical validation



- **Regression testing** to validate the performance after updates made to the algorithm
- **Testing should cover:**
 - Existing use cases that were tested using the algorithm before the updates were made
 - New use cases that are part of the updates and changes
- Changes/updates will need to be **redeployed** to reflect it for the users
- Redeployment testing and validation will depend on location of software:
 - **Cloud infrastructure**
 - **Device**
- Updates and redeployment of the algorithm should be documented as per regulatory requirements

Handling updates and changes: Clinical validation



- Clinical validation should certify that the validity of the data and clinical relevance has not changed
- Any significant modification to the algorithm or software must undergo validation by a clinician.
- The categorization of updates as minor or major is determined by the project team that would consist of a clinician
- Utilize Medical Device Regulators Forum (**IMDRF**) risk categorization principles to determine if the update falls under major/minor
- The patient safety and validity of the result is the primary importance when the algorithms are updated
- The vendors maintaining the software and the algorithm is accountable and responsible for the validity of the biomarker results provided by the tool

CASE STUDY

1

Objective: Develop a predictive biomarker using software/technology on the Welltory COVID-19 dataset

Focus: Understand the steps involved in biomarker creation

Data Insights: Even distribution of positive and negative labeled data enhances performance

VERIFICATION AND VALIDATION

2

Analytical validation

Confirm the accuracy, reliability, and precision of the software



Clinical validation

Assess biomarker's clinical relevance and predictive value

Provided recommendations for validation process

IMPACTS OF UPDATES & CHANGES

3

Explore how changes affect analytical and clinical validation process



Consider implications for biomarker performance and reliability

Thank You!



Better Health, Brighter Future

A thick, solid red vertical bar located on the left side of the slide.

Questions?