

17th
Annual
International

STC 2017

Software Testing Conference

December 07 - 08, 2017
Bangalore

Organised by

QAI

In association with

ETI

AI and Software Testing

Vikram Raghuwanshi
Senior Consultant
Capgemini Technology Services





This whitepaper gives an introduction to Artificial Intelligence; how AI touches our lives on a day to day basis, & how AI systems work.

It discusses aspects of the application of AI in QA and explains what it takes to develop an AI based solution.

The paper showcases a Capgemini AI based tool and concludes by highlighting the benefits of AI in QA.



Jeff Bezos, Amazon.com Inc

“Artificial Intelligence (AI) is a self improving, enabling horizontal layer that is solving problems that were in the realm of science fiction for the past several decades”



FORBES

Artificial Intelligence is the broader concept of machines being able to carry out tasks in a way that we would consider “smart”.



SAS

Artificial intelligence (AI) makes it possible for machines to learn from experience, adjust to new inputs and perform human-like tasks.



STANFORD

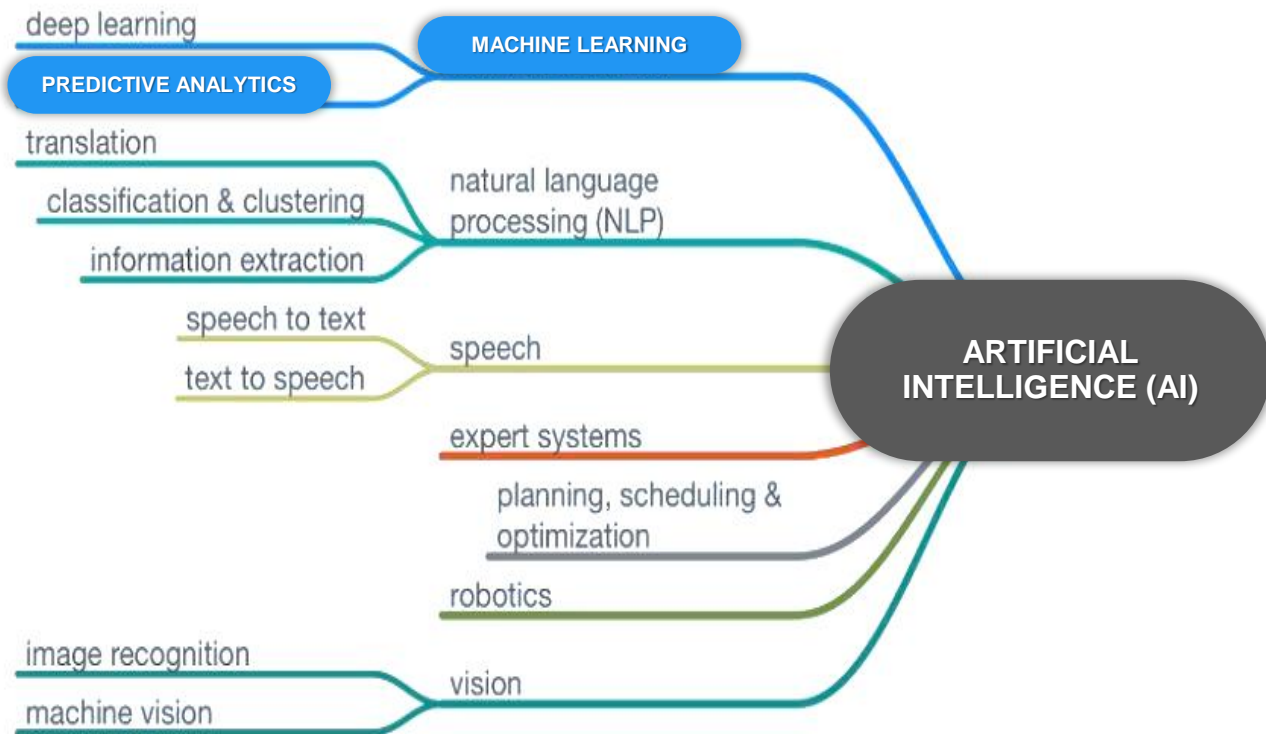
It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.



Artificial Intelligence is the broader concept of enabling machines to carry out tasks and to think and reason just as humans do

ML is a current application of AI based on the idea that machines be given access to data and allowed to learn for themselves

Data analytics, forms the core of “learning process” that machines undergo to make sense of the data that it digests





BANKING

Fraud Detection

- AI has the ability to identify fraudulent behavior, as well as identify what the next pattern of suspicious behavior will be
- AI algorithms can preempt such fraudulent transactions and can lead to huge costs savings for banks and its customers

Meeting regulatory requirements

- AI is used to ensure that regulatory requirements are met and that data is kept with monitoring done on a real-time basis. This allows issues to be flagged a lot sooner

Boost customer engagement

- Artificial intelligence will assist in the creation of customized and intelligent products and services, with new features, more intuitive interactions (e.g. speech) and advisory skills (e.g. personal financial management)



HEALTHCARE

Computer aided diagnosis

- AI is being used extensively to read and interpret complex radiology, pathology reports to help doctors arrive at early diagnosis

High risk groups identification

- AI is being fed huge volumes of data related to patients historical medical records that helps in identifying whether a patient is in a high risk group for any particular disease say stroke, cardiovascular diseases or cancer.

Epidemic outbreak prediction

- ML and AI technologies are also being applied to monitoring and predicting epidemic outbreaks around the world, based on data collected from satellites, historical information on the web, real-time social media updates, and other sources.
- Support vector machines and artificial neural networks have been used, for example, to predict malaria outbreaks, taking into account data such as temperature, average monthly rainfall, total number of positive cases, and other data points.

How AI based systems work



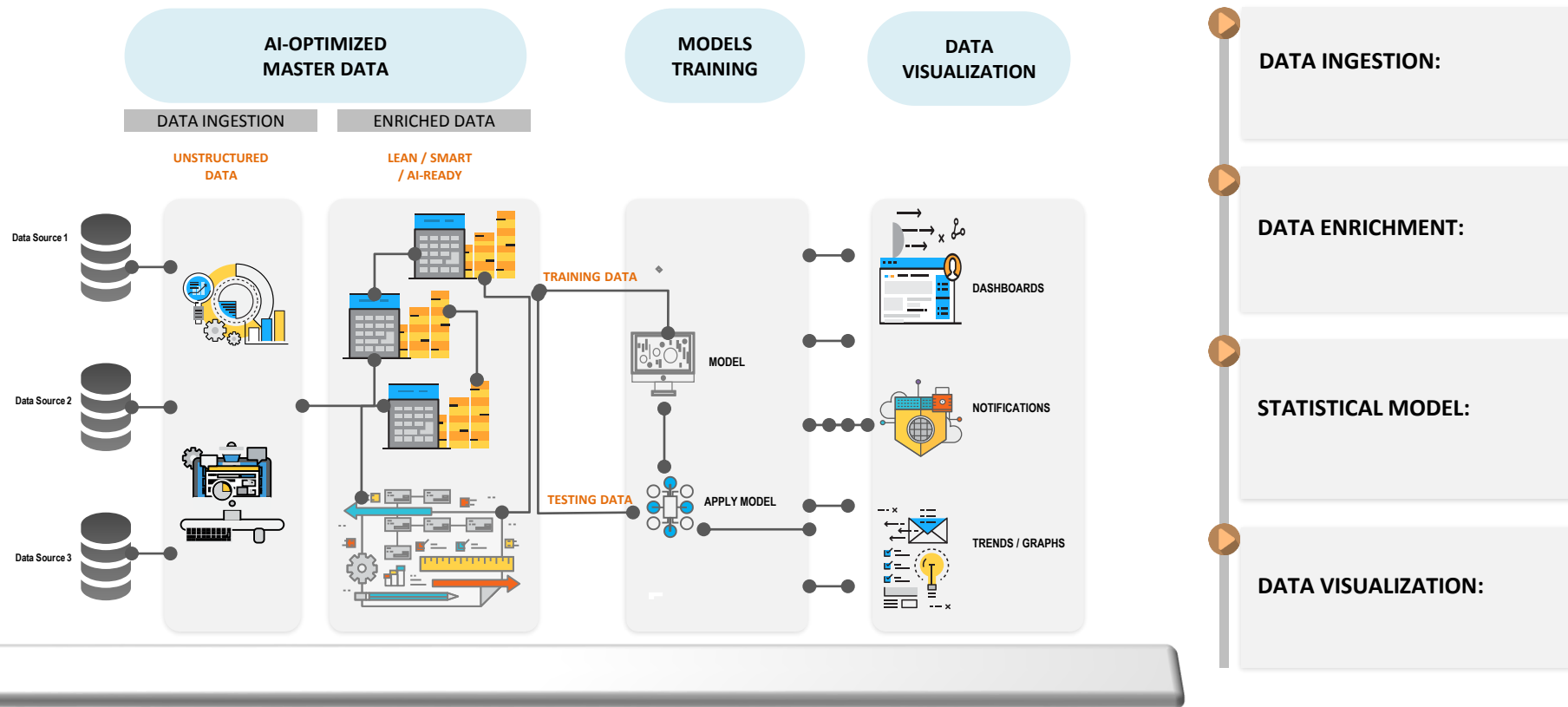
AI based systems employ a host of advanced statistical algorithms that processes huge volumes of input data in order to come up with extremely useful insights and/or predictions

An AI workflow typically involves 3 major parts:

Model: the system that makes predictions

Parameters: used by model to form its decisions

Learner: the system that adjusts the parameters, and in turn the model, by looking at differences in predictions versus actual outcome.



What you need to build AI systems (Languages)

Python

- Python is a popular and widely used language and a rising star for machine learning
- It can seamlessly be used with the data structures and other frequently used AI algorithms

- Choice of Python for AI projects also stems from the fact that there are plenty of useful libraries that can be used in AI for example, **Numpy** offers scientific computation capability, **Scipy** for advanced computing and **Pybrain** for machine learning in Python
- Easy to learn, program and debug
- Tons of reading materials and helpful resources available online.

R

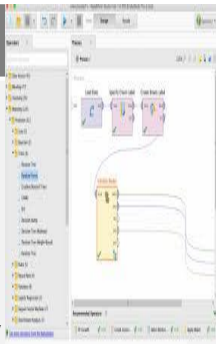
- R is a workhorse for statistical analysis and by extension **machine learning**
- Hugely popular among data scientists, statisticians, data analysts and enthusiasts
- It is the default language to understand and explore data using statistical methods and graphs

- Encompasses enormous number of **machine learning algorithms**, and advanced implementations too written by the developers of the algorithm
- Complex statistical algorithms can be written in a few lines of code
- Plenty of extremely useful R packages such as **dplyr**, **plyr** to manipulate data; **ggvis**, **lattice** and **ggplot2** to visualize data; **caret** for machine learning

What you need to build AI systems (Platforms & Skillsets)

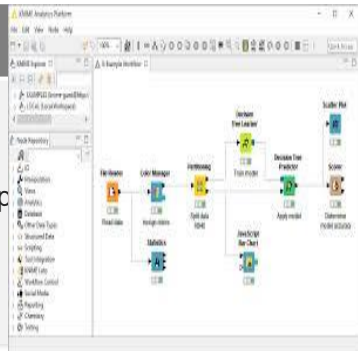
RapidMiner

- Licensed but offers a free version with restriction on input data rows
- Graphical user interface-based data science platform
- Ease of use for all data scientists and skill levels & has “large and vibrant” user community



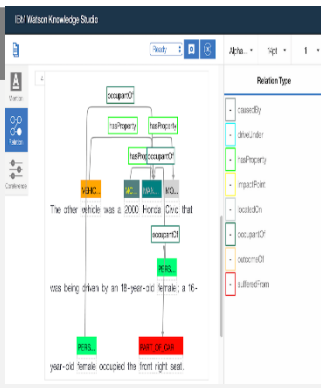
KNIME

- Opensource data analysis platform
- Flexible, and low total cost of ownership
- Graphical user interface-based data science platform



IBM Watson

- Advanced and a powerful analytics platform with a breath of offerings
 - Knowledge Studio
 - Natural Language Classifier
 - Tone analyzer
 - Speech to Text, Text to Speech
- Possess advanced NLP features
- High costs involved



Skillsets

- **Data Scientists:** Competent in using advanced Machine Learning & Statistical algorithms
- **Big Data Architect:** Well versed with handling huge datasets running into terabytes

Applications of AI in Software Testing

Test Plan Support

Test Suite Analysis

RTM analysis

Identification of right
Test Cases

Continuous Integration



An futuristic AI application that will address various aspects of testing

Chat Bots

Defect analysis

Exhaustive Reporting

Preserving Of test
Artifacts

Machine Learning

- **Data Acquisition & Cleansing**

- Test plan
- Test Cases
- Defects
- Requirements Traceability Matrix
- Code Quality Reports
- Downtime Logs

- **Relevant access to data source**

- AI Engine needs access rights to various systems. For example, for ALM it needs query editor access, & developer access for TFS / SVN etc.,

- **Access to Data Science Platform**

- RapidMiner, KNIME, RStudio

- **Resources skilled in Data Science, Machine Learning, AI**

AI solution gets various data feeds; it then gives critical predictive analysis to the testing teams across project's lifecycle assisting them to keep adjusting to various moving parameters influencing outcomes of the project.

Asset analytics:

- Captures key metrics and make recommendations on improving leverage from assets.
- Analyses asset's usage trend, execution duration, attached criticality through requirements, defects mapping etc.

Staff analytics:

- Assists in understanding the distribution of capabilities, specific skills among the complete team & to track real time progress,
- Predicts which resource is best suited for what tasks (& auto assigns such tasks to identified resources based on bandwidth)

CUSTOMER ANALYTICS:

- Sentiment analysis by analyzing customer feedback from multiple social channels
- Identifies (from customer feedback) key issues in applications designs, behavior, usage patterns etc. allowing testing teams to plan better around digital testing coverage.

RIGHTSIZING SOLUTIONS:

- Combination of various analytical utilities guiding testing organization to size their testing appropriately real time

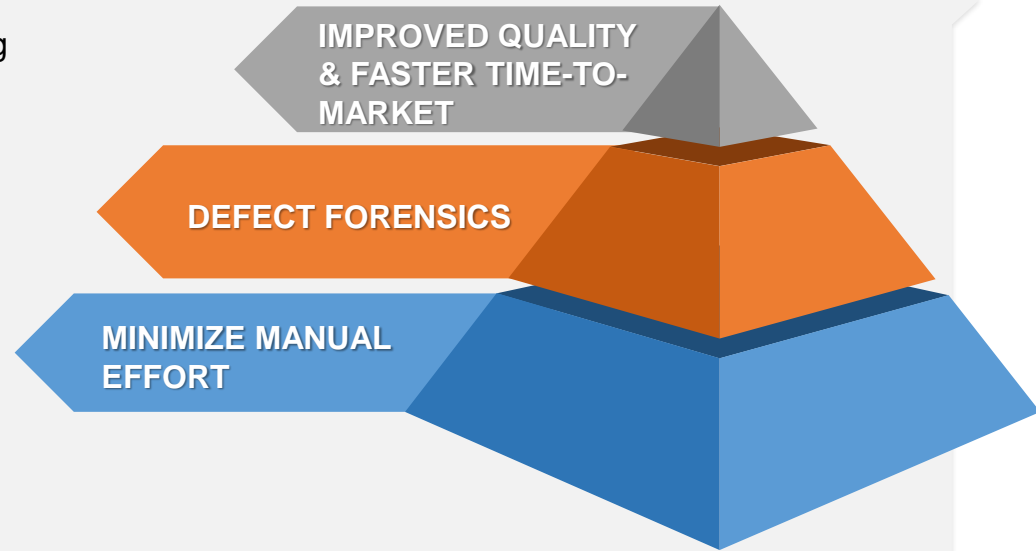
DELIVERY ANALYTICS:

- When to start / stop Testing
- Provides real time test data consumption and business requirement coverage

PREDICTIVE ANALYTICS:

- Predicting likelihood of a Test Case failure (Defect Prediction)
- Predicting vulnerable (prone to defects) modules in the Application Under Test
- Predicting the right Test Case pack for a new requirement

- Shift left testing for faster time-to-market & improved quality
- AI application will effectively help in generating and optimizing test cases, prioritizing testing and automation, enhancing UI testing and reducing tedious analysis tasks.
- Unattended execution – In case of failures during Test Execution, AI system shall execute recovery routines and achieve unaided execution
- Scientific approach for defect forensics: (Application Stability, defect hotspots, Defect pattern and failure predictions)
- Social media analytics: Mechanism to analyze data from social media to build trends by demographics; forming the basis of test suites



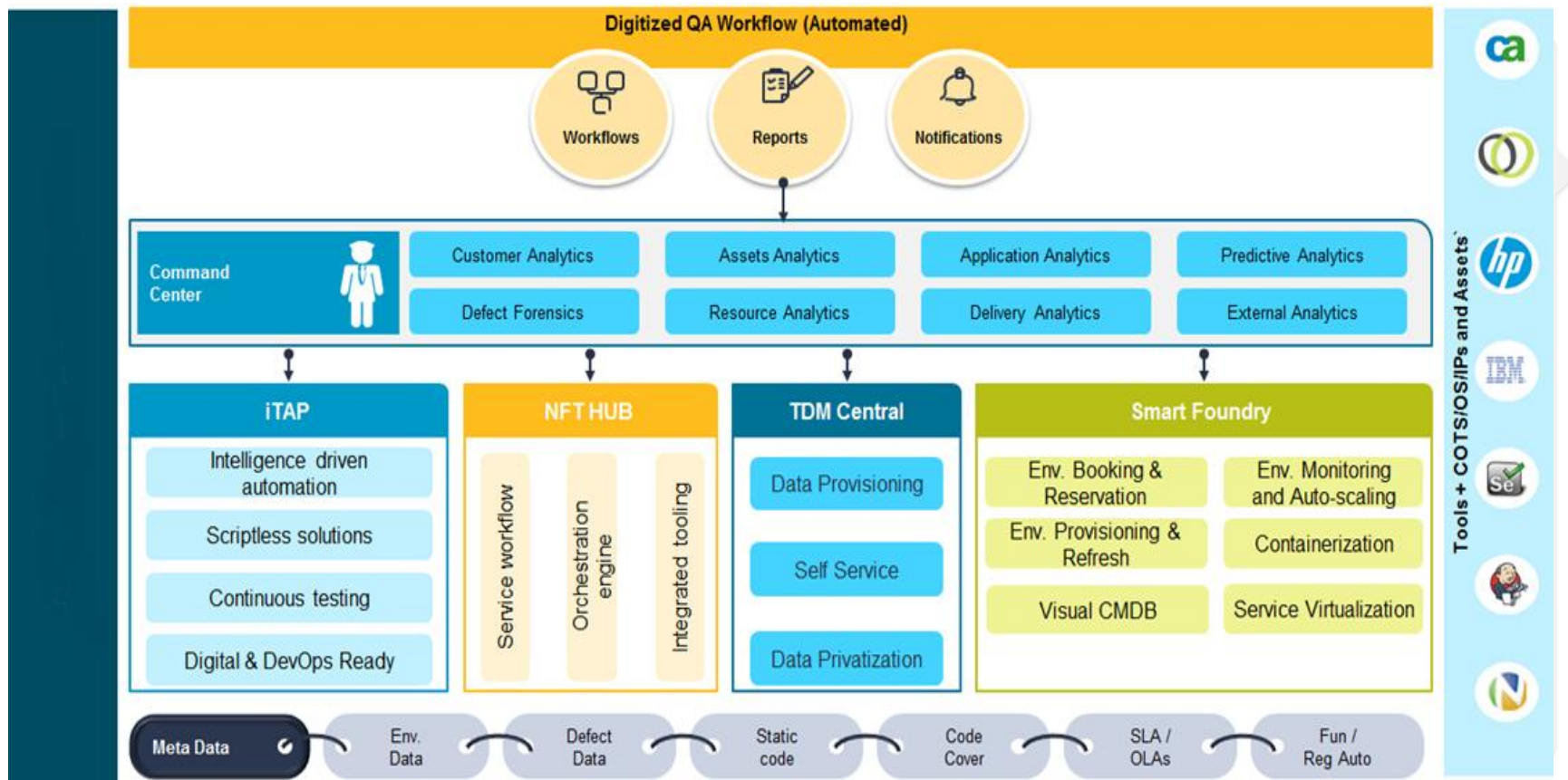
Example of an actual AI enabled System for QA

Capgemini's Smart QA is an analytics-driven platform that helps banks, financial institutions, and insurers build an end-to-end ecosystem, which includes testing assets, a test environment, test data and performance data

It has An analytics engine, which includes cognitive capabilities powered by smart techniques and insights which test architects can use to prioritize what needs to be tested, optimize testing efforts, and identify areas of improvement.

Smart QA also knows how much testing is required and which assets should be used.

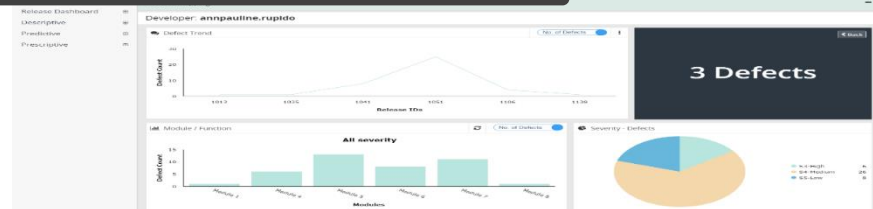




Resource Prediction

- Dashboard aggregates data of defects that have been logged by the developers
- The aggregated historical defect data will be analyzed based on severity and type of defect and prediction can happen for future developer defects during releases
- The dashboard offers analysis, retrospection and prescription of defects that have been logged

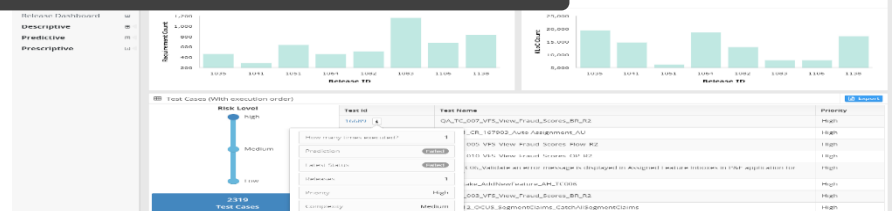
Resource Prediction



Risk Based Testing

- Exact test cases that need to be executed based on priority.
- Prediction for next test execution
- Complexity of test case execution

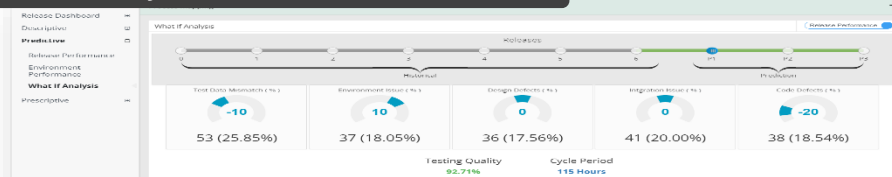
Risk Based Test Case Selection



What-If analysis

- What if Analysis predicts defects at code & release level
- What if Analysis offers various levers which can help predict percentage defects that may happen
- The levers can be adjusted as per desired test quality and release cycle period.

What-If analysis



TEST CASE RECOMMENDATION

List of Available Test Cases

Selected Test Case TC3 for execution

Test Case Design Studio

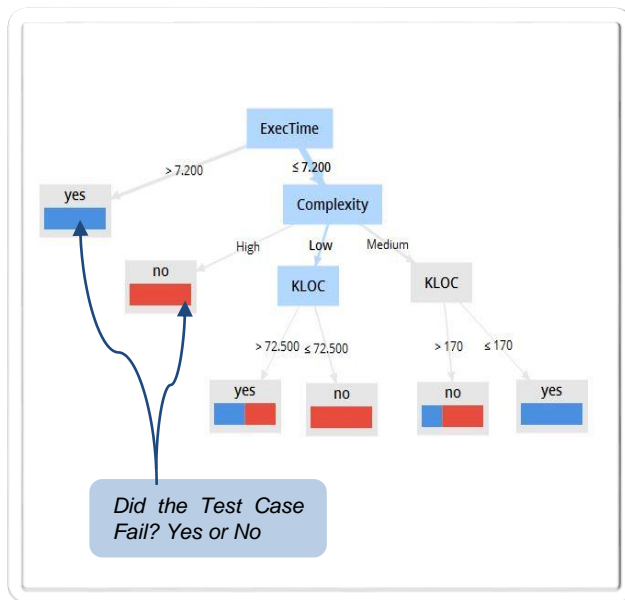
Test Case Recommendation

Recommended Test Cases list for TC3

"Support" & "Conf"

This Use case is on similar lines as the product recommendation on popular e-commerce sites

AUTO / SELF ROOT CAUSE ANALYSIS



The Decision Tree algorithm works on the concept of Entropy and looks to classify the input data based on gaining maximum information and stability.

PREDICT THE TEST CASE FAILURE

Test Case	Actual TC status		Predicted TC status	
	Failed		prediction	confidence
TC1	yes		yes	0.991
TC2	yes		yes	1.000
TC3	no		yes	0.778
TC4	no		no	0.001
TC5	no		no	0.022
TC6	yes		yes	0.882
TC7	yes		no	0.067
TC8	no		no	0.001

The Logistic Regression's output helps in assigning a quantitative value to each Test case



AI and its applications are increasingly gaining acceptance in every vertical and/or service line that the organizations operate in.

Industry gurus believe testing is certainly one area which could immensely benefit from AI.

At this juncture it is worthwhile to note that while AI is promising & emerging, it is fairly complex and has a steep learning curve initially.

However with proper planning on resources, use cases, software, hardware and on overall architecture, one shall definitely be able to overcome the said challenges and leverage the benefits for a successful AI based solution for QA

- Wikipedia - en.wikipedia.org
- https://en.wikipedia.org/wiki/Evolutionary_computation
- <https://www.ibm.com/watson/>
- <https://www.openai.com>
- <http://www.ipsoft.com>
- <http://www.wipro.com/holmes/>
- <https://www.capgemini.com/service/capgemini-smart-qa-for-financial-services/>
- <https://www.knime.com/knime-analytics-platform>
- Fausett, L. Fundamentals of Neural Networks: Architecture,
- Algorithms and Applications. Prentice Hall 1994.



Vikram is having 12+ years of rich IT experience.

He has worked on support, S/W Development, S/W testing and Automated testing.

He has keen interest in current & emerging technologies. His educational qualifications include BSc. (PCM) & MCA



THANK YOU