ISACA®
Trust in, and value from, information systems
San Diego Chapter

ISACA San Diego Chapter

August 17, 2017

Conducted by Bill Bonney, VP Product Marketing and Chief Strategist

HOOSH®
PROTECTING AND POWERING YOUR DATA®

Introduction to Auditing Artificial Intelligence

Bill Bonney

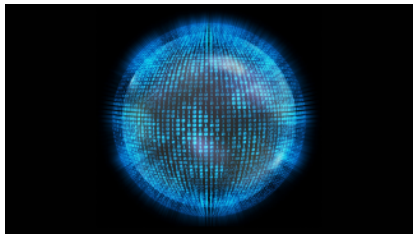
HOOSH®
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

IBM 1401 – We put a man on the Moon

Thinking Machines

What is Artificial Intelligence?



Intel Xeon +FPGA -- Intel Saffron memory based reasoning system – Intel Xeon E5 Family of processors

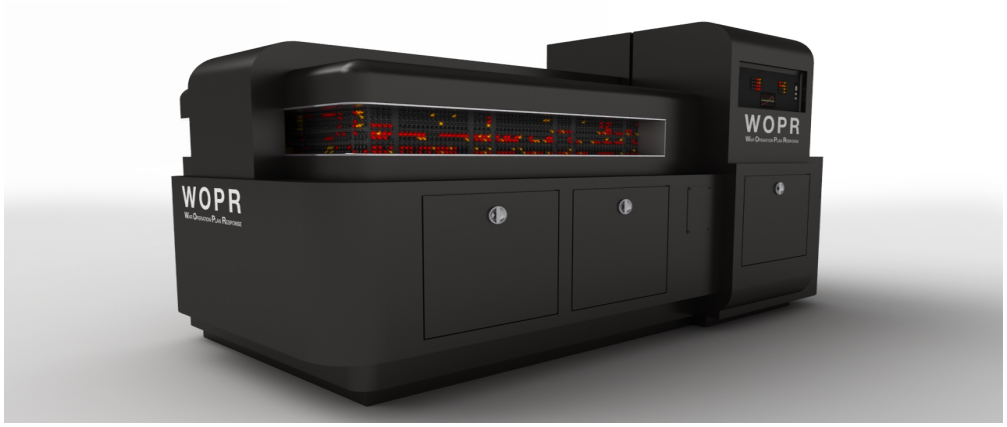
HOOSH®
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

Entertainment Icons

- **Movie Stars: Colossus & HAL 9000**
- **TV Stars: The Bat Computer (Burroughs to Dell PowerEdge) & Lt. Commander Data (Positronic brain)**
- **Book Characters: Mycroft Holmes (Mike) & Ship**
- **Comic Book Characters: Braniac and Cerebro**

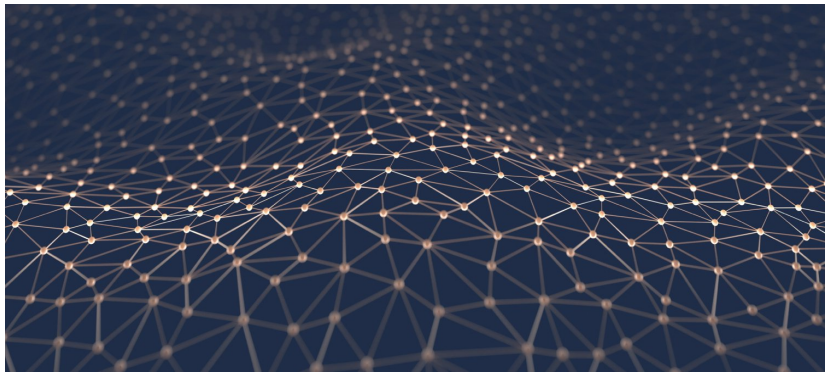
Well over 600 references in a list of fictional appearances by AI "beings."



Do You Want To Play A Game?

HOOSH
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved



“Artificial Intelligence is the study of how to make computers do things at which, at the moment, people are better.”

–Elaine Rich, 1983

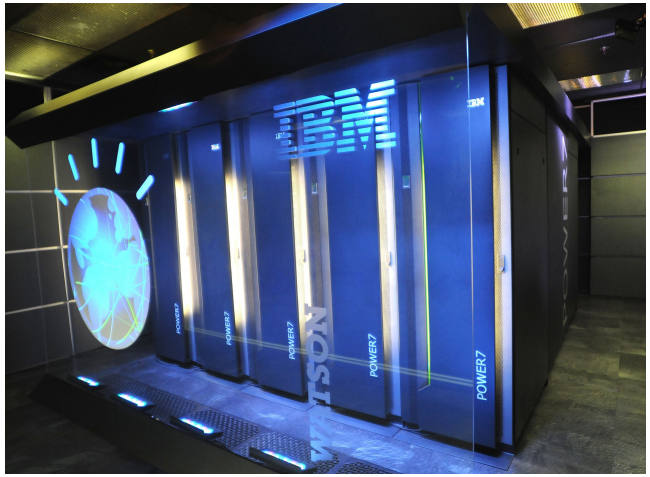
The CMU professor, not the Dynasty actress.

HOOSH®
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

Alan Turing (1912-1954):

"The Imitation Game" referred not to breaking the Enigma machine but to answering the question: Can machines think?



Or more precisely, can a machine imitate a human so well that you can't tell the difference between a human and a machine?

HOOSH
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

Brief History of AI

- The field of artificial intelligence research was founded at a workshop held on the campus of Dartmouth College during the summer of 1956. Minsky & McCarthy.

- Over promising and under achieving led the U.S. and British governments to cut funding in 1973, ushering in the first AI winter 74-80.

- In the 80's visionary leadership shifted to Japan. Investment money flowed, but the hardware and software was still not up to the task. After $billions invested, money dried up, and the second AI winter set in, 87-93.

- Not everyone gave up. IBM and others continued to invest and eventually, Deep Blue started making progress and a chess program was created and a challenge issued to Gary Kasparov in 1996. But, Deep Blue was defeated. Was winter coming again?

HOOSH®
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

Do you think life is one big game...

- **1997 IBM's Deep Blue defeats the chess champ Gary Kasparov (10 ** 120 positions (8X8 board))**
  - And he was NOT happy
- **2011 Watson beats Jennings and Rutter on Jeopardy**
  - Itchy trigger finger aside
- **2016 Google's DeepMind's AlphaGo beat Go Master Lee Se-Dol (10 ** 761 positions (19X19 board)**
- **2016 Machines faced off in a special Defcon CTF!**

1080 particles in the observable universe!**

Deep Blue, at the Computer History Museum

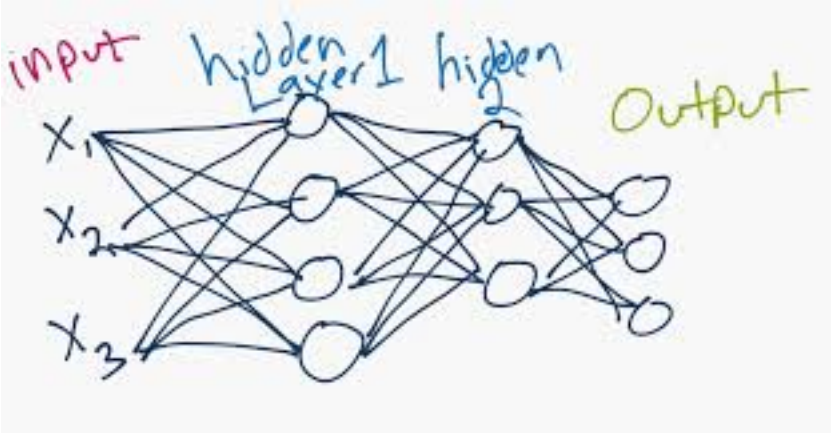
HOOSH
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

Rise of the Machine: Learning

Narrow and general types of AI

- **Narrow (or "weak")AI:**
  - Non sentient
  - Focused on one task



- **General (or "strong") AI:**
  - Sentient
  - Can apply intelligence to any task

HOOSH
PROTECTING AND POWERING YOUR DATA®
Copyright(c) Bill Bonney. 2017 - All Rights Reserved

Fear of AI: Hawkins, Gates, Musk

There are valid causes of the fear of AI:

Speed: Machines can act much faster than humans in certain tasks. Think: Programmed Trading.

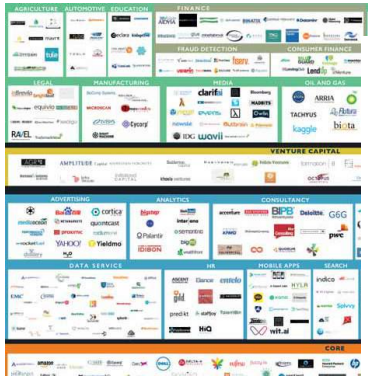
Empathy: Machines will evolve differently than humans and will not have the same instincts of social justice.

Dependence: As our world becomes more complex, we rely on computers more and more. Will we lose control?

HOOSH®
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

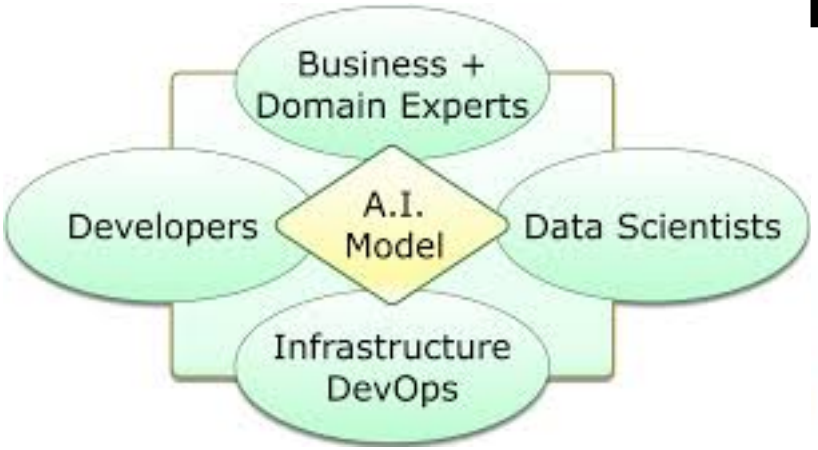
How are Artificial Intelligence and Machine Learning useful in business?



AI in Financial services

- In the field of Finance, artificial intelligence has long been used. Some applications of Artificial Intelligence are:–
- Credit authorization screening
- Mortgage risk assessment
- Project management and bidding strategy
- Financial and economic forecasting
- Risk rating of exchange-traded, fixed income investments
- Detection of regularities in security price movements
- Prediction of default and bankruptcy
- Security/and or Asset Portfolio Management

How might we audit such a thing?



HOOSH
PROTECTING AND POWERING YOUR DATA®

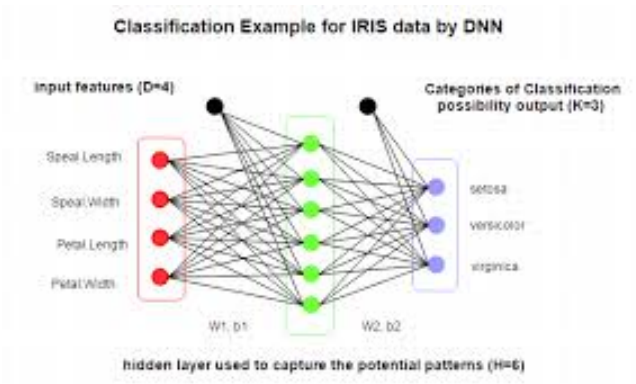
Copyright(c) Bill Bonney. 2017 - All Rights Reserved

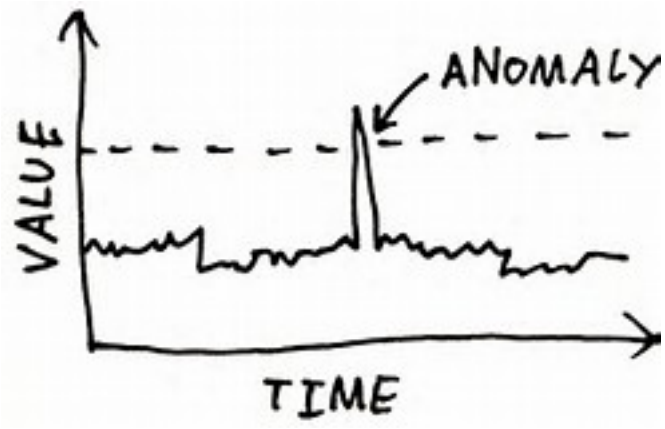
Machine Learning vs Predictive Analysis

The principal differences between the two approaches are:

- **Predictive analysis** is limited to predicting futures based on historical data, is use-driven requiring static instructions to re-run output models to improve result accuracy; whereas

- **Machine learning** covers a large variety of problems, is data driven and updates its models dynamically and automatically with no human intervention.





Copyright(c) Bill Bonney. 2017 - All Rights Reserved

Machine Learning: Practical Definitions

Machine learning offers a set of tools, mathematical models and algorithms, that can predict outcomes from large data sets but then evolve and adapt its results as and when new data is added. The power of this approach is that it can achieve outcomes based on priorities which become increasingly reliable over time as it consumes more data to feed the models it is continually enhancing. Real-time decisions can then be made and acted upon without human intervention.

Predictive analytics is usually described as being a subset of machine learning which itself is seen as one aspect, albeit from an enterprise point of view the most valuable at present, of artificial intelligence.

Machine learning is a blanket term for non-human intervention during the processing and analysis of data, there are several different tasks associated with machine learning.

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

Machine Learning: Practical Definitions

- **Supervised learning**: this is the most common approach used and describes the use of sample or test outcome data to assist a machine in learning how to infer a mapping function from a set of given input data. The machine can then independently use the mapping function with unseen input data, making learning decisions and inferences as it progresses.

  In other words, the machine is given a helping hand in understanding the task before it and, once it has learnt its lesson, can be left to its own devices thereafter. This category of supervised learning can be further sub-divided into a set of regression (the output is a value) and classification (the output is a category) problems.

- **Semi-supervised learning**: occurs when part of the input data set is labelled but the rest (usually the majority) is not – hence learning inferences must be made based on only partial help. Consequently, the problem is a hybrid of supervised and unsupervised learning and is very common where data labelling is difficult or resource intensive and so expensive or inefficient.

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

Machine Learning: Practical Definitions

Unsupervised learning: unlike supervised learning the machine is not given any output data to calibrate its mapping function and is responsible for generating a set of results based on its own interpretation of the underlying structure, relationships or distribution of the input data. The most common approaches taken to achieve this set of tasks are:

- **Clustering** – creating clusters of data derived from patterns found in the input data
- **Association or attribute selection** – creating rules that apply to large sets of input data
- **Anomaly detection** – identifying events or activities that do not fit the expected pattern of behavior. Even here, there are further sub-categories: supervised (first labeling what is considered 'normal' data), **unsupervised** (letting the machine determine what is normal and what falls outside that) and **semi-supervised** (using test data to identify what is a normal and then applying to the input data).

HOOSH®
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

Uses and Applications of Machine Learning in the Enterprise

Activity	Application	Description
Sales & Marketing	Understand your markets	Understand the optimal market and customer segments for your products by geography, demographic, delivery and time-to-buy, leading to more direct targeting.
	Campaign management	The above task would enable a machine to learn the best approaches to cross-channel marketing as well as cross-selling and up-selling.
	Know Your Customer	Get to know your customers better from their lifetime purchasing history, complaint or service resolution as well as an analysis of any social media feedback delivering insights and patterns of behavior.
	Pricing	A deeper understanding of your markets enables a machine learning task to derive a more nuanced approach to customer-specific pricing. Or, depending on your industry/products, the capacity to generate dynamic pricing based on real-time events.
	Forecasting	Machine learning derived from any or all of the above tasks has the potential to forecast likelihood to buy, building up an overall quarterly product demand analysis and inventory plan into the warehouse.
Security	Attack prevention	One of the most prevalent applications of machine learning is in anticipating potential cyber-attacks based on unusual behavior and/or anomaly detection from a wide number of sensors, alerts, alarms and other forms of event notifications.
	Fraud prevention	Any business (or individual) using online payments is open to fraud. Given the complexity and heterogeneity of input data it is required to assess, there are not resources available to detect and stop fraudsters, making machine learning' usability to detect patterns of behavior and adapt to variations a must for every business concerned about being compromised. Several different machine learning methods and techniques are available to address the problem depending on context.
Operations	Supply optimization	From analyzing network utilization, power usage and heat distribution whether in a data center, a restaurant or a factory floor, machine learning can make real-time, event-based optimization and regulation supply and demand decisions for these resources.
Finance	Risk management	A crucial aspect of many businesses, particularly in the banking and finance communities, is to be able to make an accurate and near real-time assessment of an individual or a business's credit worthiness.

HOOSH®
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

Examples of Machine Learning by Industry

Industry	Application	Description
Healthcare	Proactive health management	From wearable devices to hospital bedside and theatre monitoring systems generating real-time alerts and outputs combined with a patient's recent and historical medical history, a machine can learn to provide an up to the minute diagnosis of disease identification, risk assessment and even suggested treatments and medicines
Manufacturing	Process optimization	Most manufacturing processes involve a complex variety of factors and variables that can easily involve wastage or lead to mistakes. In most cases traditional quality control methods cannot respond or be processed as quickly as required – or as effectively as a machine can learn and interpret.
Smart Cities	Traffic management	Congestion management based upon historical and real-time pattern analysis can improve traffic flow as well as enable the prioritization of emergency services with controlled disruption.
	Driverless cars	In addition to helping the flow of traffic through cities, machine learning is going to be vital in enabling driverless cars to communicate with each other and the plethora of road-side sensors.

HOOSH®
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

Hershey's Twizzlers

Hershey installed 22 IoT sensors in cooking vats in one candy manufacturing facility

- Sensors took measurements every second, creating over 600,000 readings per 8-hour production shift.
- In roughly 90 days, they had 60 million data points
- Fed the data to machine learning algorithm provided by Microsoft Machine Learning Studio on Azure

They had no data scientists on staff and did the project to see if they could democratize ML to the entire IT function.

The training consisted of a few simple steps.

- They first divided the data into two sets
  - one for learning
  - one for validating that learning was complete.
- The data was then fed into the application and outliers were identified.
- The outliers were confirmed or removed
- Then they trained the algorithm to learn what "good" and "bad" look like in the readings.

This was completed using roughly two-thirds of the data points

- The remaining third was used to validate that the training was successful.
- The correlation engine was allowed to run against the data and find the data points that would be helpful in predicting successful outcomes.

It turned out that four of the 22 sensors provided predictive value.

- Once training was accomplished, the equipment could take adjustments directly from the application.
- The adjustments were made after the prediction engine ran against data collected in 15-minutes intervals.
- Capture 20,000 raw total readings or 3,600 predictive readings every 15 minutes of sensor readings
- Small adjustments were made to the controls of the holding vats.

Once fully engaged, they saved over $500,000 per year by reducing waste and overage in just one candy line at just one facility.

HOOSH®
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

Sample Audit Checklist

Completed	Review Step	Auditor
	All data collection points (including sensor data, transaction data, etc.) have been identified and validated	
	Full data set has been identified	
	Machine learning algorithms have been identified and validated	
	For new applications, the full data set has been appropriated divided into “teaching” and “validating” segments	
	Teaching results indicate teaching data set successfully input	
	Validation demonstrated that the learning objectives were met	
	Data outliers were identified	
	The outliers were confirmed or removed	
	The algorithm was trained to recognize “good” and “bad” results	
	Correlation was appropriately run on the data set and correlating data points were identified	
	Adjustment thresholds were identified tested for effectiveness and adjusted as needed	
	Results of execution, including pre- and post-adjustment, have been recorded and appropriate lessons learned captured	
	Project objectives (cost, time, error rate, etc.) were met	

HOOSH®
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

General Audit Points for Consideration

Design Evaluation

Goal – *to reduce the inherent complexity in analyzing input from a wide range of different data sources to making real-time or near real-time decisions*

The first step therefore is to assess:

- What data you will be working with
- What insights or outcomes you hope to derive from them
- How you anticipate feeding those results into the decision-making process
- The return on investment.

Executive Buy-In

Have appropriate budget and resources been allocated to allow the project to be successful.

- Just because the Hershey project proved you don't have to have an army of data scientists for each machine learning project doesn't mean that all machine learning projects can be staffed with inexperienced personnel.
- Do you have a dedicated and supported hardware/software/ networking platform to support the appropriate machine learning problem to be addressed.

HOOSH®
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved

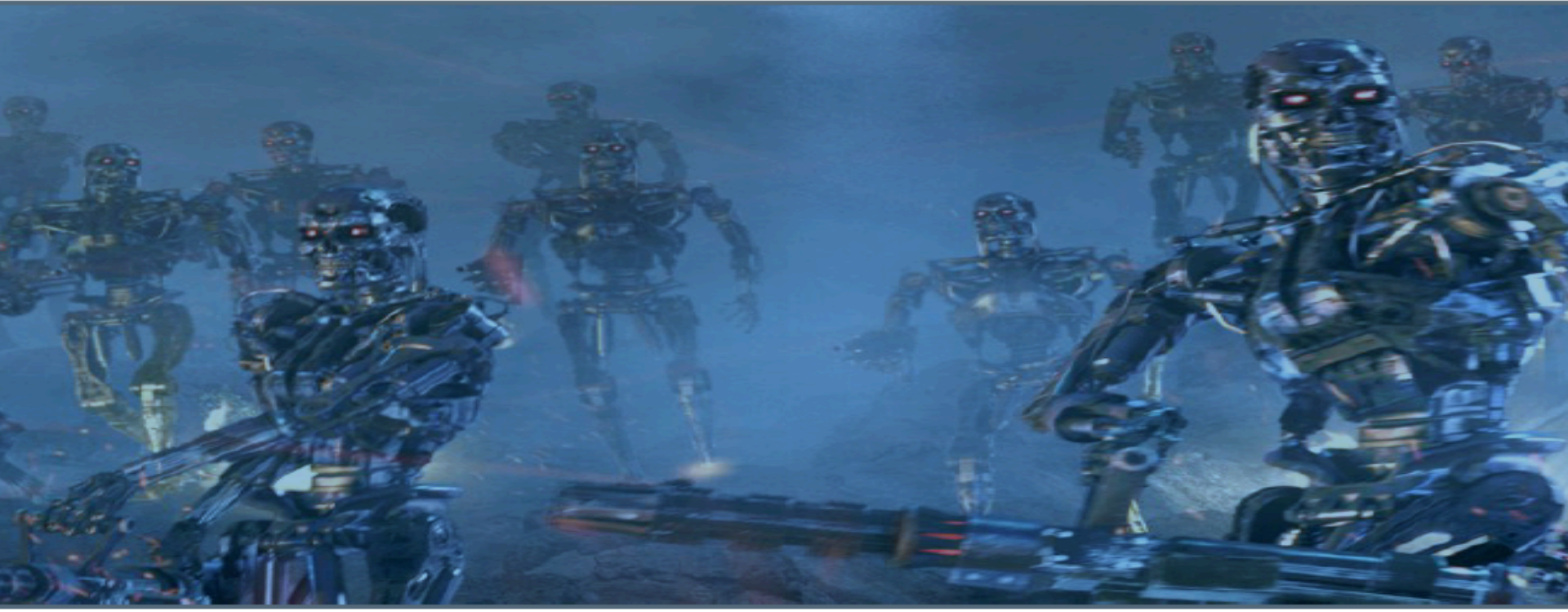
Machine Learning Packages

Here are several packages you can use to learn how to use machine learning techniques. Even a rudimentary knowledge of the actual performance of ML will help you assess the effectiveness of ML techniques in your organization.

- **Microsoft Machine Learning Studio** makes machine learning part of their cloud-based analytics package called the Cortana Intelligence Suite.
- **Amazon** offers a collection of artificial intelligence products, including Lex (natural language understanding, and automatic speech recognition), Rekognition (for image recognition), Polly (for text to lifelike speech) and machine learning to build predictive models out of existing data sets.
- **Google Cloud Prediction API** provides the ability to predict email veracity, use browsing and searching habits to predict purchases and determine what a given person might spend per day based on their spending history. Google Cloud Prediction API can integrate with App Engine, and the RESTful API is available through libraries for many popular languages, such as Python, JavaScript and .NET. The Prediction API also provides pattern-matching and machine learning capabilities.
- **IBM** has made Watson Analytics available as an API library that allows you to access the Watson IoT platform via RESTful calls. The source code is available via GitHub & is easily examined to allow teams to learn how to fully exploit the capabilities.
- **Algorithms.io** provides a cloud-hosted service to collect data, generate classification models and score new data. Random forest, support vector machine, K-Means, decision tree, logistic regression and neural network algorithms are provided.
- **FICO Analytic Cloud** offers Analytic Modeler for R, a robust computing platform to develop descriptive and predictive models using the vast statistical libraries of open source R, Powered by RStudio, and an analytic modeler scorecard to predict the likelihood of various business events, such as fraud, attrition, or propensity to buy.

IBM, Microsoft and Amazon offer (throttled) free accounts and Microsoft, Amazon and FICO provide algorithm development and solution building tools.

Copyright(c) Bill Bonney. 2017 - All Rights Reserved



Questions & Discussion

Bill Bonney
Vice President and Chief Strategy Officer
FHOOSH, Inc.
bill@fhoosh.com
@wqbonney
https://www.linkedin.com/in/BillBonney

Thank You

FHOOSH
PROTECTING AND POWERING YOUR DATA®

Copyright(c) Bill Bonney. 2017 - All Rights Reserved