JUNE 8, 2017

Machine Learning Case Studies/Use Cases

Insurance, Pharma, Healthcare

Adrian Bowles, PhD

Founder, STORM Insights, Inc.

Lead Analyst, AI, Aragon Research

info@storminsights.com

STORM INSIGHTS INC.

DATAVERSITY®

Aragon Research

Agenda

- ☑ Foundations of Industry-Specific ML Applications
- ☑ Insurance
  - ☑ Applications
  - ☑ Example(s)
- ☑ Pharma
  - ☑ Applications
  - ☑ Example(s)
- ☑ Healthcare
  - ☑ Applications
  - ☑ Example(s)
- ☑ Transforming Industries with Machine Learning

STORM INSIGHTS INC.

Foundations of Industry-Specific ML Applications

What Do Insurance, Pharma, & Healthcare have in Common?

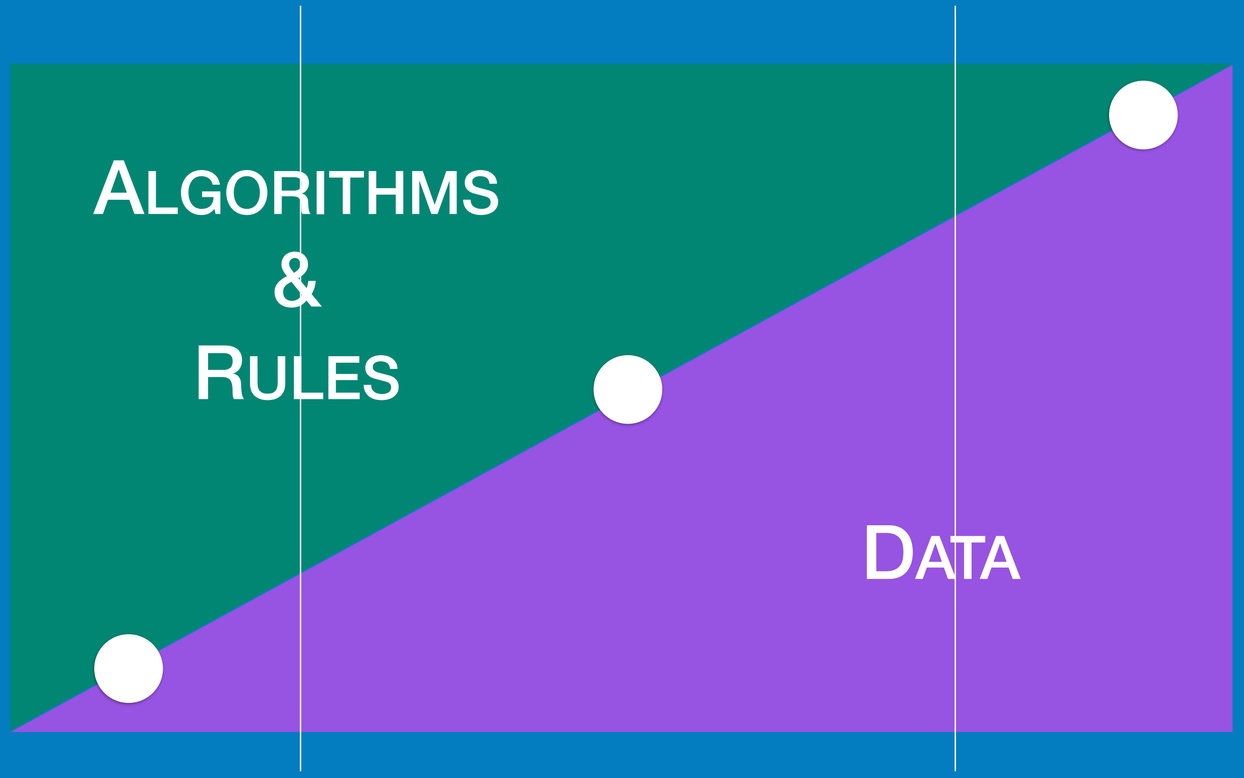
Data - Lots of It
- Historical
  - Highly Structured - Customer databases, etc.
  - Deep Structure - Journals, case notes, audio/video intake records, etc.
- Access to Streaming Data
  - Telematics, Weather, Biometrics, News

Well-defined vocabularies, data models, taxonomies
- Ranging from insurance claim codes to biochemistry models

Regulatory Issues

STORM INSIGHTS INC.

AI Learning Trends



More Data + Faster HW make
Deep Learning Practical

Deep Learning Success With Recognition
Spurs Investment

Investment Leads to Investigation
Broaden the Scope of Applications

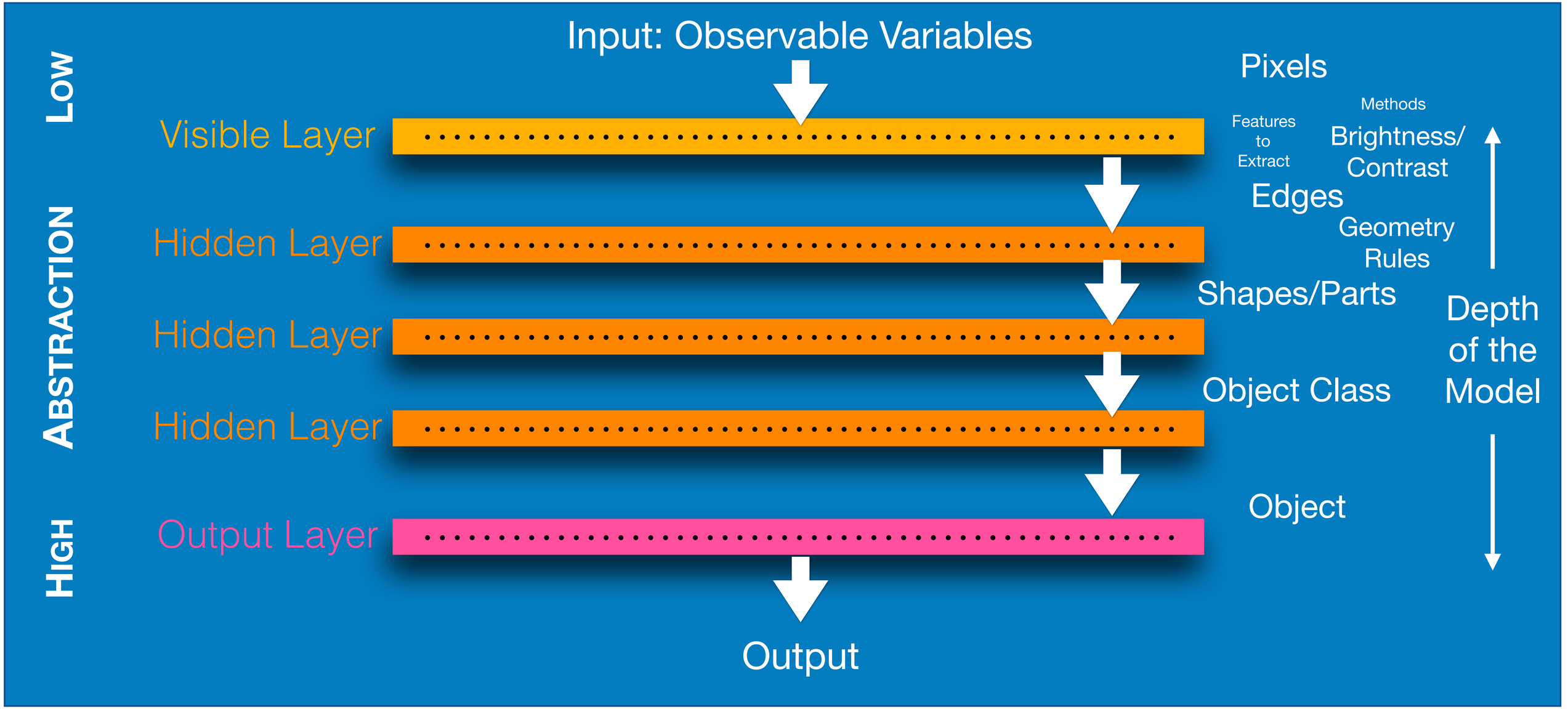
Caution for Applications Where
Transparency is Critical

New “Explainability” Research Emerges

Hybrid Solutions to Augment Intelligence
Will Thrive for Critical Applications

STORM INSIGHTS INC.

Deep Learning



Insurance Industry: Major Use Cases

Auto, Property & Casualty, Life…

- ☑ Pricing
- ☑ Risk Management
- ☑ Claim Loss Prediction
- ☑ Claim Loss Prevention
- ☑ Customer Experience
- ☑ Fraud Detection

Data

- ☑ Customer
  - Demographics and historical record
- ☑ Property
  - Actual property data, local/regional data
- ☑ Streaming
  - Near real time personal state and behavioral data ranging from social sentiment analysis to biometrics, and related news and weather

STORM INSIGHTS INC.

AXA Insurance Case Study: Auto Insurance Pricing POC

Historical Data:
7-10% of their insured drivers cause an accident annually
1% of accidents result in payments > $10,000

Challenge:
Improve over current methods to predict high risk policy holders.

Google Cloud Platform

Why Google | Products | Solutions | Launcher | Pricing | Customers | Documentation | Support | Partners

GOOGLE CLOUD BIG DATA AND MACHINE LEARNING BLOG
Innovation in data processing and machine learning technology

Using machine learning for insurance pricing optimization

Wednesday, March 29, 2017

By Kaz Sato, Staff Developer Advocate, Google Cloud

AXA, the large global insurance company, has used machine learning in a POC to optimize pricing by predicting "large-loss" traffic accidents with 78% accuracy.

The TensorFlow machine-learning framework has been open source since just 2015, but in that relatively short time, its ecosystem has exploded in size, with more than 8,000 open source projects using its libraries to date. This increasing interest is also reflected by its growing role in all kinds of image-processing applications (with examples including skin cancer detection, diagnosis of diabetic eye disease and even sorting cucumbers), as well as natural-language processing ones such as language translation.

We're also starting to see TensorFlow used to improve predictive data analytics for mainstream business use cases, such as price optimization. For example, in this post, I'll describe why AXA, a large, global insurance company, built a POC using TensorFlow as a managed service on Google Cloud Machine Learning Engine for predicting "large-loss" car accidents involving its clients.

Understanding the use case

Approximately 7-10% of AXA's customers cause a car accident every year. Most of them are small accidents involving insurance payments in the hundreds or thousands of dollars, but about 1% are so-called large-loss cases that require payouts over $10,000. As you might expect, it's important for AXA adjusters to understand which clients are at higher risk for such cases in order to optimize the pricing of its policies.

Toward that goal, AXA's R&D team in Japan has been researching the use of machine learning to predict if a driver may cause a large-loss case during the insurance period. Initially, the team had been focusing on a traditional machine-learning technique called Random Forest. Random Forest is a popular algorithm that uses multiple Decision Trees (such as possible reasons why a driver would cause a large-loss accident) for predictive modeling. Although Random Forest can be effective for certain applications, in AXA's case, its prediction accuracy of less than 40% was inadequate.

In contrast, after developing an experimental deep learning (neural-network) model using TensorFlow via Cloud Machine Learning Engine, the team achieved 78% accuracy in its predictions. This improvement could give AXA a significant advantage for optimizing insurance cost and pricing, in addition to the possibility of creating new insurance services such as real-time pricing at point of sale. AXA is still at the early stages with this approach — architecting neural nets to make them transparent and easy to debug will take further development — but it's a great demonstration of the promise of leveraging these breakthroughs.

STORM INSIGHTS INC.

AXA INSURANCE: AUTO INSURANCE PRICING: NEW RISK MODELING APPROACH

Identified ~70 risk factors
Age,
Address,
Vehicle type,
Previous Accidents,
Original Channel,
Car Age Range
…

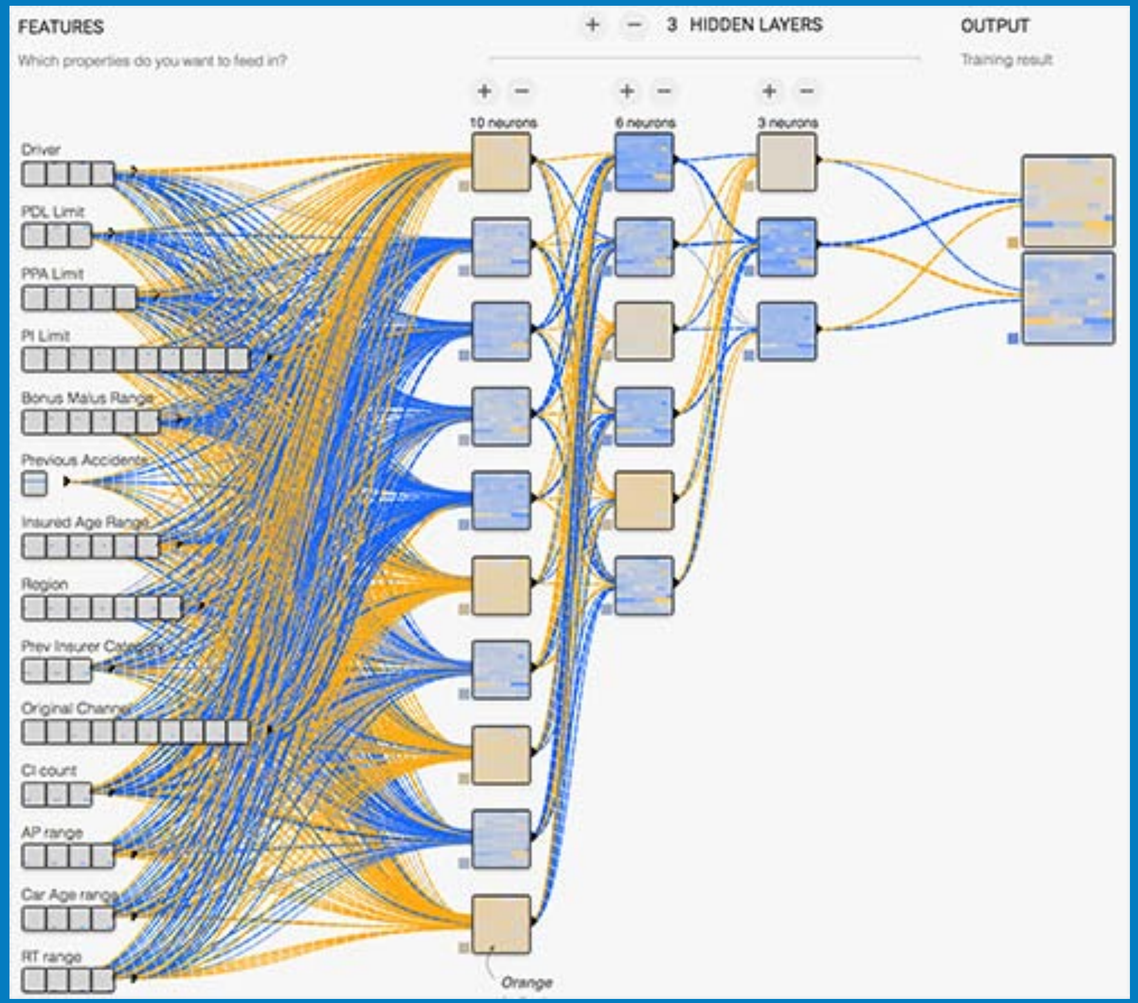
STORM INSIGHTS INC.

AXA INSURANCE: AUTO INSURANCE PRICING: TECHNICAL APPROACH

Using TensorFlow on the
Google Cloud
Machine Learning Engine

Input Layer

Age,
Address,
Vehicle type,
Previous Accidents,
Original Channel,
Car Age Range
…



Source: Google Cloud Platform Blog

STORM INSIGHTS INC.

AXA INSURANCE: AUTO INSURANCE PRICING: RESULTS

Predicted Risk Accurately ~78%

enabling cost optimization and new services

STORM INSIGHTS INC.

Pharma

Major Use Cases

- ☐ Drug discovery
- ☐ Clinical Trials
- ☐ Analysis (Biochemical)

Data

- ☑ Biochemistry
- ☑ Clinical Trial Case Data
- ☑ Journals, News…

STORM INSIGHTS INC.

PHARMA: DRUG DISCOVERY & CLINICAL TRIALS

Pharmaceutical Research and Manufacturers of America®
950 F Street, NW Suite 300, Washington, DC 20004

PhRMA
RESEARCH • PROGRESS • HOPE

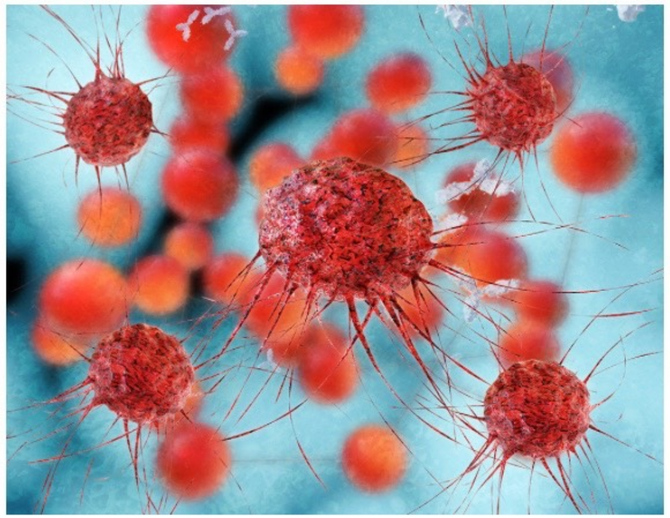
OUR NETWORK | MEMBERS SITE | PRESS ROOM | EVENTS | STAY CONNECTED

ABOUT | ADVOCACY | SCIENCE | PATIENTS | RESOURCES | BLOG

REPORT

Medicines in Development for Immuno-Oncology 2017 Report

PhRMA | June 1, 2017 | SHARE THIS



More than 240 immuno-oncology treatments in development are accelerating progress for patients with cancer. The rapid pace of scientific advances has helped usher in a new era of medicine for cancer patients over the last decade. Biopharmaceutical researchers' understanding of the underlying biological mechanisms that initiate and control cancer cell growth have created promising new avenues for treatment advances. Research into the role of the body's immune system in fighting cancer has yielded some of the most exciting new advances, resulting in a new wave of immunotherapies specifically targeting cancers.

DOWNLOAD

STORM INSIGHTS INC.

PHARMA: OPTIMIZING VIRTUAL SCREENING WITH MACHINE LEARNING

Machine learning based search space optimisation for drug discovery

Sign In or Purchase to View Full Text

3 Paper Citations

123 Full Text Views

Related Articles

Profile-based string kernels for remote homology detection and motif extraction

Prediction of RNA-Binding Residues in Protein Sequences Using Support Vector Mac...

Adaptive thresholding for multi-label SVM classification with application to pro...

View All

3 Author(s)

Upul Senanayake ; Rahal Prabuddha ; Roshan Ragel

View All Authors

Abstract | Authors | Figures | References | Citations | Keywords | Metrics | Media

Abstract:
Drug discovery research has progressed to a place where it essentially counts on high performance computer systems and huge databases for its victory. As such, Virtual Screening (VS), a computationally intensive process, plays a major role in the systematic drug designing process for pressing diseases. Therefore, it is imperative that the VS process has to be made as fast as possible in order to efficiently dock the ligands from huge databases to a selected protein receptor, targeting a drug. The extremely high rate of increase of the number of ligands in the databases makes it impossible to tackle this problem only by improving the computing resources. Therefore, researchers work on an orthogonal technique, where they use soft computing to reduce the search space through identifying the ligands that are non-dockable, hence improving the throughput as a whole. Machine Learning (ML) can be used to train a binary classifier that can classify the ligands into two known classes: dockable and non-dockable ligands. In this paper, for the first time, we use three ML techniques (Support Vector Machines, Artificial Neural Networks and Random Forest) on a single problem domain (a Protease receptor of HIV) and evaluate the performance rendered by the respective models. We show that such classification improves the throughput by two folds with around 90% accuracy. In addition, we propose and use a technique for constructing a training set to be used for ML in VS applications in the instance of a non-synthesised receptor.

Published in: Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 2013 IEEE Symposium on

Date of Conference: 16-19 April 2013

INSPEC Accession Number: 13752071

Date Added to IEEE *Xplore*: 12 September 2013

DOI: 10.1109/CIBCB.2013.6595390

ISBN Information:

Publisher: IEEE

STORM INSIGHTS INC.

PHARMA: A NEW ML APPROACH USING 3-D PROTEIN MODELING

THE VARSITY

The University of Toronto's Student Newspaper Since 1880

Work with us!

NEWS COMMENT FEATURES ARTS SCIENCE SPORTS PHOTO VIDEO MAGAZINE

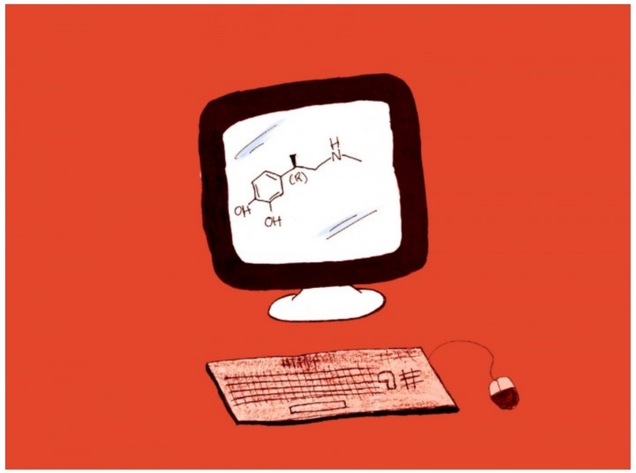
Accelerating drug discovery with machine learning

U of T researchers develop machine learning algorithms to accelerate drug discoveries

By Utkarshna Sinha

Published: 11:12 am, 27 February 2017

under **Science**



OREOLUWA/THE VARSITY

http://var.st/26q

Researchers at U of T have developed new machine learning algorithms to determine the 3D structures of proteins, paving the way for faster and more efficient drug discoveries and diagnostic methods.

In order to be effective, drugs must bind to specific proteins in a cell in the right orientation. They do this by changing the conformation of a protein, which results in a change in that protein's function. Knowing the 3D structure of a protein can significantly enhance understanding of how they work in the body and consequently, aid in the development of drugs targeting the potential harmful effects of these proteins at unprecedented speeds and efficacies.

"The ability to discover 3D structures of protein molecules is one of the major goals of the field of structural biology. Proteins, which are the building blocks of every biological process, are tiny molecular machines that interact, bind, move and work together to make life happen," said Ali Punjani, a PhD student who is working on these algorithms under Professor David Fleet, Chair of the Computer and Mathematical Sciences department at UTSC.

This team of researchers, along with Dr. Marcus Brubaker, an Assistant Professor at York University who worked on these algorithms with Fleet as a postdoctoral researcher, have developed new machine learning algorithms to dramatically speed up the process of solving the 3D structure of a protein.

This novel approach does not require a scientist to guess what the protein could look like.

"We employ an algorithm for finding 3D structures that effectively

Related Articles

Mapping Protein Interactions To Tackle Cancer

U of T's Stagljar Lab has built a protein interactome map to improve cancer treatments

UTM To Get New Centre For Medicinal Chemistry

A $7 million donation from Orlando Corporation will fund the project

Accelerating Entrepreneurship

The Hatchery hosts its Accelerator Weekend in collaboration with Y Combinator

Study On Social Behaviours Of Mice Reveals Potential Autism Cause

New U of T research has linked autism to a protein

STORM INSIGHTS INC.

PHARMA: A NEW ML APPROACH

Stanford | News

Search Stanford news...

Home Find Stories For Journalists Contact

APRIL 3, 2017

Stanford researchers create deep learning algorithm that could boost drug development

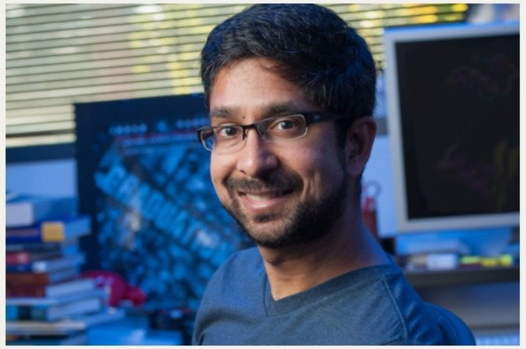
Combining computer science and chemistry, researchers show how an advanced form of machine learning that works off small amounts of data can be used to solve problems in drug discovery.

BY TAYLOR KUBOTA

Artificially intelligent algorithms can learn to identify amazingly subtle information, enabling them to distinguish between people in photos or to screen medical images as well as a doctor. But in most cases their ability to perform such feats relies on training that involves thousands to trillions of data points. This means artificial intelligence doesn't work all that well in situations where there is very little data, such as drug development.

Vijay Pande, professor of chemistry at Stanford University, and his students thought that a fairly new kind of deep learning, called one-shot learning, that requires only a small number of data points might be a solution to that low-data problem.

"We're trying to use machine learning, especially deep learning, for the early stage of drug design," said Pande. "The issue is, once you have thousands of examples in drug design, you probably already have a successful drug."



Stanford chemistry Professor Vijay Pande and his students see a future for machine learning in the early stages of drug development. *(Image credit: L.A. Cicero)*

The group admitted the idea of applying one-shot learning to drug design problems was farfetched – the data was likely too limited. However, they'd had success in the past with machine learning methods requiring only hundreds of data points, and they had data available to test the one-shot approach. It seemed worth a try.

Much to their surprise, their results, published April 3 in *ACS Central Science*, show that one-shot learning methods have potential as a helpful tool for drug development and other areas of chemistry research.

STORM INSIGHTS INC.

PHARMA: IBM PATENTS DRUG DISCOVERY MACHINE LEARNING MODELS

IBM

News room > News releases >

News room

- **News releases**
- Press kits
- Image gallery
- Biographies
- About IBM
- News room feeds
- Global news rooms
- News room search
- Media contacts

Related links
- IT Analyst support center
- Investor relations

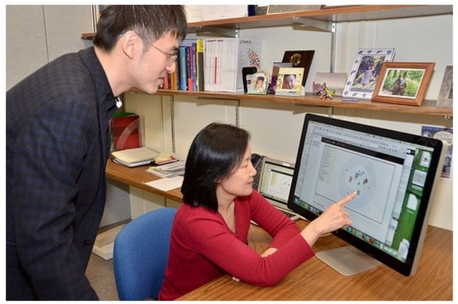
IBM Patents Machine Learning Models for Drug Discovery

Invention could accelerate discovery of more effective and safer drugs

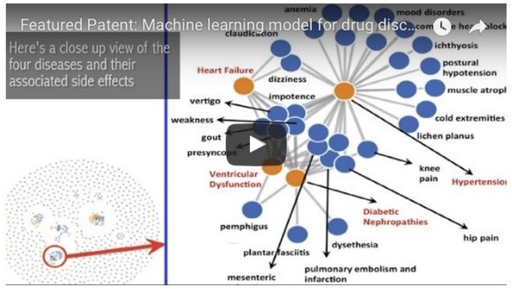
Select a topic or year

- News release
- Related XML feeds
- Contact(s) information
- Related resources

Armonk, N.Y. - 07 Apr 2017: IBM (NYSE: IBM) today announced that its scientists have been granted a patent on machine learning models to predict therapeutic indications and side effects from various drug information sources. IBM Research has implemented a cognitive association engine to identify significant linkages between predicted therapeutic indications and side effects, and a visual analytics system to support the interactive exploration of these associations.



IBM Research staff member Ping Zhang (left) and Program Director, Center for Computational Health Jianying Hu (right) with their newly patented invention to help drug discovery researchers identify which drug indications are typically linked to which side effects.



This approach could help researchers in pharmaceutical companies to generate hypotheses for drug discovery. For instance, strongly correlated disease-side-effect pairs identified by the patented invention could be beneficial for drug discovery in many ways. One could use the side-effect information to repurpose existing treatments (e.g. drugs causing postural hypotension could be potential candidates for treating hypertension). If a new drug is being designed for a disease that is strongly correlated with severe side effects, then special attention could be paid to controlling the formulation and dosing of the drug in the clinical trials to prevent serious safety issues.

IBM was granted U.S. Patent 9,536,194: **Method and system for exploring the associations between drug side-effects and therapeutic indications** for this invention.

IBM News Room Twitter
Join the conversation

Share
Facebook
Twitter
LinkedIn

Images
Cognitive Health Patent

Engage IBM
Contact a media relations representative
Site feedback

RSS
Subscribe to our latest news releases
View more news room feeds

STORM INSIGHTS INC.

HEALTHCARE

Major Use Cases

- ☐ Epidemic Prediction & Monitoring
- ☐ Diagnosis
- ☐ Treatment

Data

- ☑ Patient

  Demographics and historical record/EHR

- ☑ Treatment/Outcome Data

  Taxonomies, drug interaction guides, journals, case records…

- ☑ Streaming

  Near real time personal state and behavioral data ranging from social sentiment analysis to biometrics, and related news and weather

STORM INSIGHTS INC.

HEALTHCARE & ML CHALLENGE

How important is an explanation to trust?

Not as important as demonstrated efficacy.

Aspirin
Lithium
Placebos

MIT Technology Review

Log in / Register Search

Topics+ Top Stories Magazine Events More+

Subscribe

Rewriting Life

Deep Learning Is a Black Box, but Health Care Won't Mind

New algorithms are able to diagnose disease as accurately as expert physicians.

by Monique Brouillette April 27, 2017

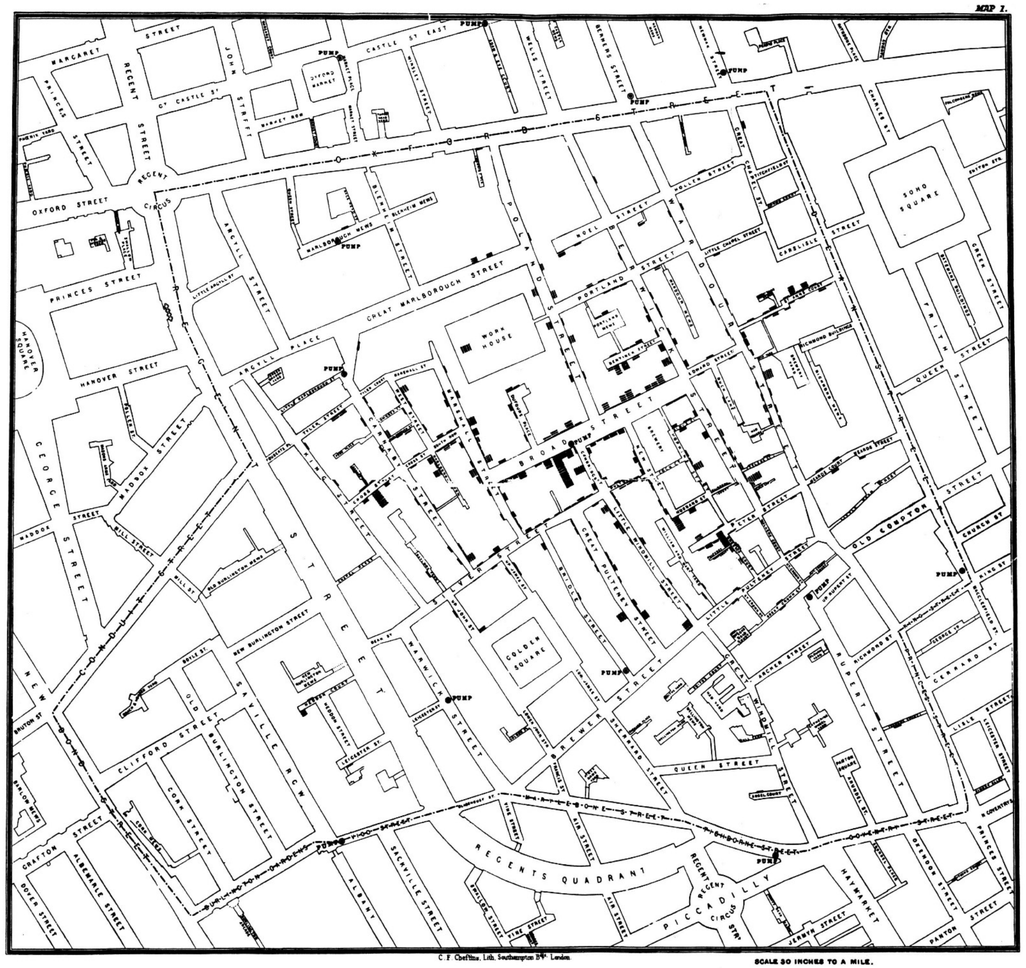
Earlier this year, artificial intelligence scientist Sebastian Thrun and colleagues at Stanford University demonstrated that a "deep learning" algorithm was capable of diagnosing potentially cancerous skin lesions as accurately as a board-certified dermatologist.

The cancer finding, reported in *Nature*, was part of a stream of reports this year offering an early glimpse into what could be a new era of "diagnosis by software," in which artificial intelligence aids doctors—or even competes with them.

Experts say medical images, like photographs, x-rays, and MRIs, are a nearly perfect match for the strengths of deep-learning software, which has in the past few years led to breakthroughs in recognizing faces and objects in pictures.

STORM INSIGHTS INC.

HEALTHCARE: INFECTIOUS DISEASE MANAGEMENT



1854 Broad Street cholera outbreak. (2017, May 26). In Wikipedia, The Free Encyclopedia. Retrieved 14:10, June 8, 2017

John Snow's Map of Cholera Death Clusters

International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 4 Issue 12, December 2015

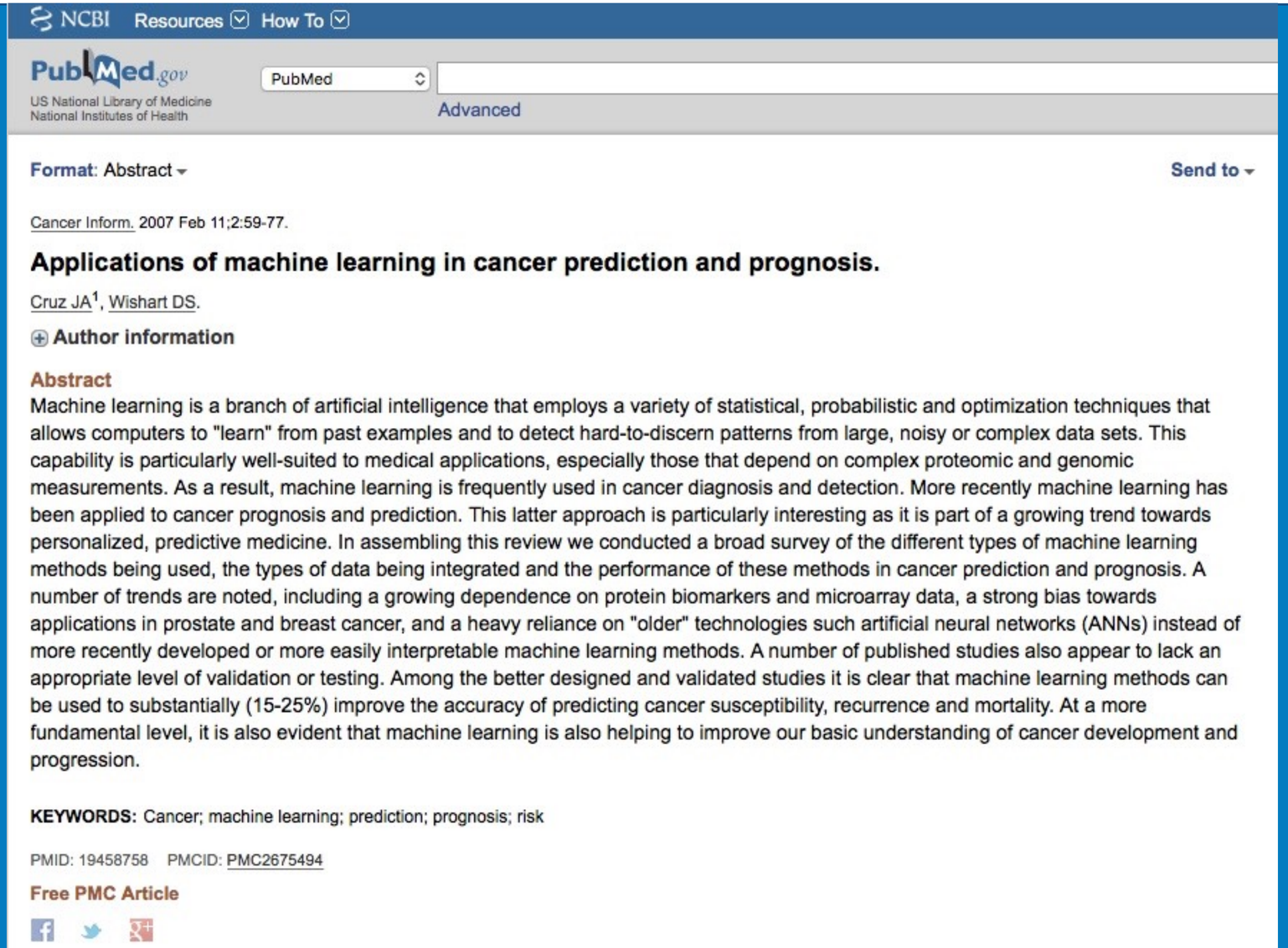
Malaria Outbreak Prediction Model Using Machine Learning

Vijeta Sharma[1],Ajai Kumar[2],Lakshmi Panat[3],Dr. Ganesh Karajkhede[4],Anuradha lele[5]

[1] *Project Engineer,Applied Artificial Intelligence Group ,C-DAC,Pune.*
[2] *Head of Department, Applied Artificial Intelligence Group, C-DAC,Pune.*
[3] *Principle Technical Officer,Applied Artificial Intelligence Group,C-DAC,Pune,*
[4] *Health Informatics Domain Expert,Centre For Development of Advanced Computing,Pune,*
[5] *Joint Director,Applied Artificial Intelligence Group,C-DAC,Pune*

STORM INSIGHTS INC.

Machine Learning for Cancer Prediction and Prognosis: Note the Date



NCBI Resources How To

PubMed.gov
US National Library of Medicine
National Institutes of Health

PubMed

Advanced

Format: Abstract

Send to

Cancer Inform. 2007 Feb 11;2:59-77.

Applications of machine learning in cancer prediction and prognosis.

Cruz JA[1], Wishart DS.

Author information

Abstract

Machine learning is a branch of artificial intelligence that employs a variety of statistical, probabilistic and optimization techniques that allows computers to "learn" from past examples and to detect hard-to-discern patterns from large, noisy or complex data sets. This capability is particularly well-suited to medical applications, especially those that depend on complex proteomic and genomic measurements. As a result, machine learning is frequently used in cancer diagnosis and detection. More recently machine learning has been applied to cancer prognosis and prediction. This latter approach is particularly interesting as it is part of a growing trend towards personalized, predictive medicine. In assembling this review we conducted a broad survey of the different types of machine learning methods being used, the types of data being integrated and the performance of these methods in cancer prediction and prognosis. A number of trends are noted, including a growing dependence on protein biomarkers and microarray data, a strong bias towards applications in prostate and breast cancer, and a heavy reliance on "older" technologies such artificial neural networks (ANNs) instead of more recently developed or more easily interpretable machine learning methods. A number of published studies also appear to lack an appropriate level of validation or testing. Among the better designed and validated studies it is clear that machine learning methods can be used to substantially (15-25%) improve the accuracy of predicting cancer susceptibility, recurrence and mortality. At a more fundamental level, it is also evident that machine learning is also helping to improve our basic understanding of cancer development and progression.

KEYWORDS: Cancer; machine learning; prediction; prognosis; risk

PMID: 19458758 PMCID: PMC2675494

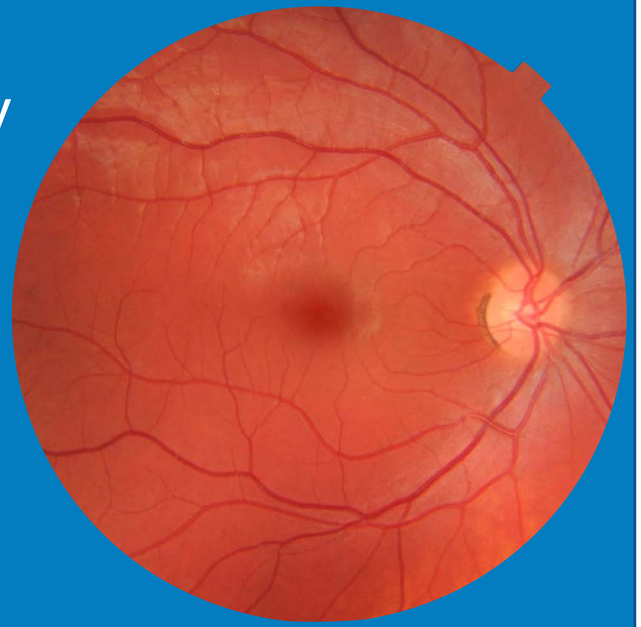
Free PMC Article

STORM INSIGHTS INC.

Healthcare Example: Diabetic Retinopathy Diagnosis

Retinal damage caused by diabetes, which may lead to blindness

- Accounts for 12% of new cases of blindness in the US annually
- Leading cause of blindness for ages 20-64
- At least 90% of new cases could be reduced with proper treatment and monitoring
- Affects up to 80%of people with diabetes for 20+ years
- Often has no early warning signs
  - Non-proliferative diabetic retinopathy (NPDR) is the first stage
    - No symptoms
    - Signs not visible to the eye
    - Patients can have 20/20 vision
    - Detected by fundus photography to see micro aneurisms

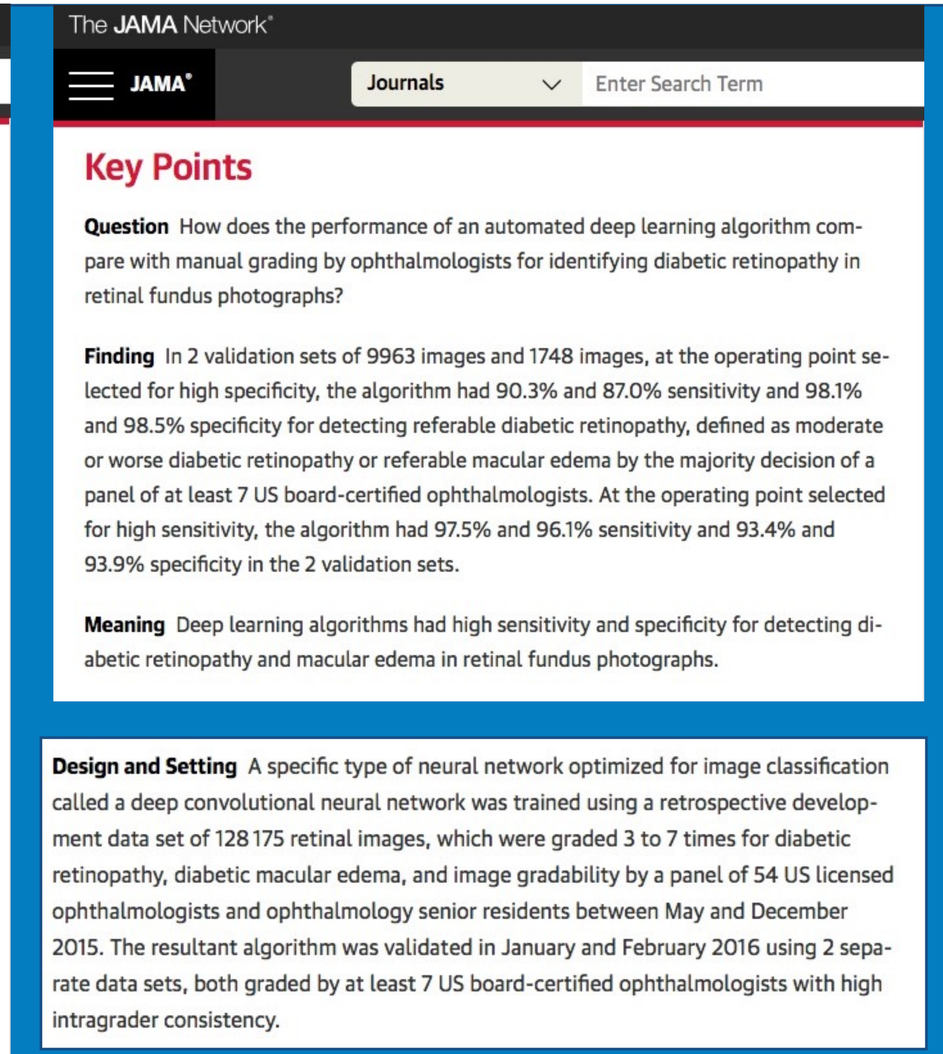


"Medical gallery of Mikael Häggström 2014". *WikiJournal of Medicine* **1** (2). DOI:10.15347/wjm/2014.008. ISSN 2002-4436. Public Domain.

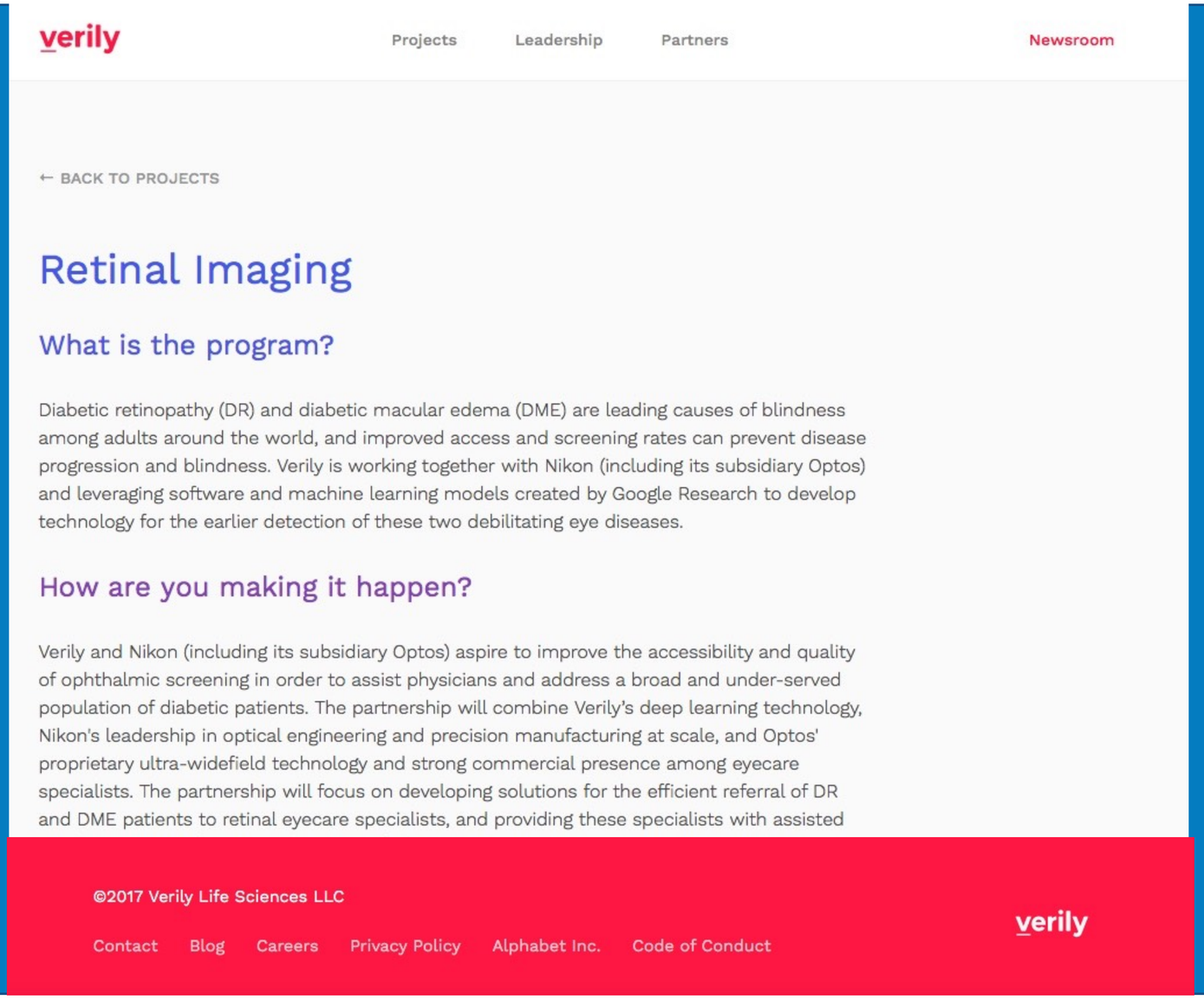
https://en.wikipedia.org/w/index.php?title=Diabetic_retinopathy&oldid=783900921

STORM INSIGHTS INC.

Diabetic Retinopathy: Google, Verily, UT Austin, UC Berkeley et al



Diabetic Retinopathy: Verily (Google/Alphabet) & Nikon



verily

Projects Leadership Partners Newsroom

← BACK TO PROJECTS

Retinal Imaging

What is the program?

Diabetic retinopathy (DR) and diabetic macular edema (DME) are leading causes of blindness among adults around the world, and improved access and screening rates can prevent disease progression and blindness. Verily is working together with Nikon (including its subsidiary Optos) and leveraging software and machine learning models created by Google Research to develop technology for the earlier detection of these two debilitating eye diseases.

How are you making it happen?

Verily and Nikon (including its subsidiary Optos) aspire to improve the accessibility and quality of ophthalmic screening in order to assist physicians and address a broad and under-served population of diabetic patients. The partnership will combine Verily's deep learning technology, Nikon's leadership in optical engineering and precision manufacturing at scale, and Optos' proprietary ultra-widefield technology and strong commercial presence among eyecare specialists. The partnership will focus on developing solutions for the efficient referral of DR and DME patients to retinal eyecare specialists, and providing these specialists with assisted

Contact Blog Careers Privacy Policy Alphabet Inc. Code of Conduct

verily

https://verily.com/projects/interventions/retinal-imaging/

Diabetic Retinopathy: Microsoft

Microsoft News Center India Our Company Our Products Press Tools Blogs &

Microsoft, L V Prasad Eye Institute and Global Experts Collaborate to Launch Microsoft Intelligent Network for Eyecare

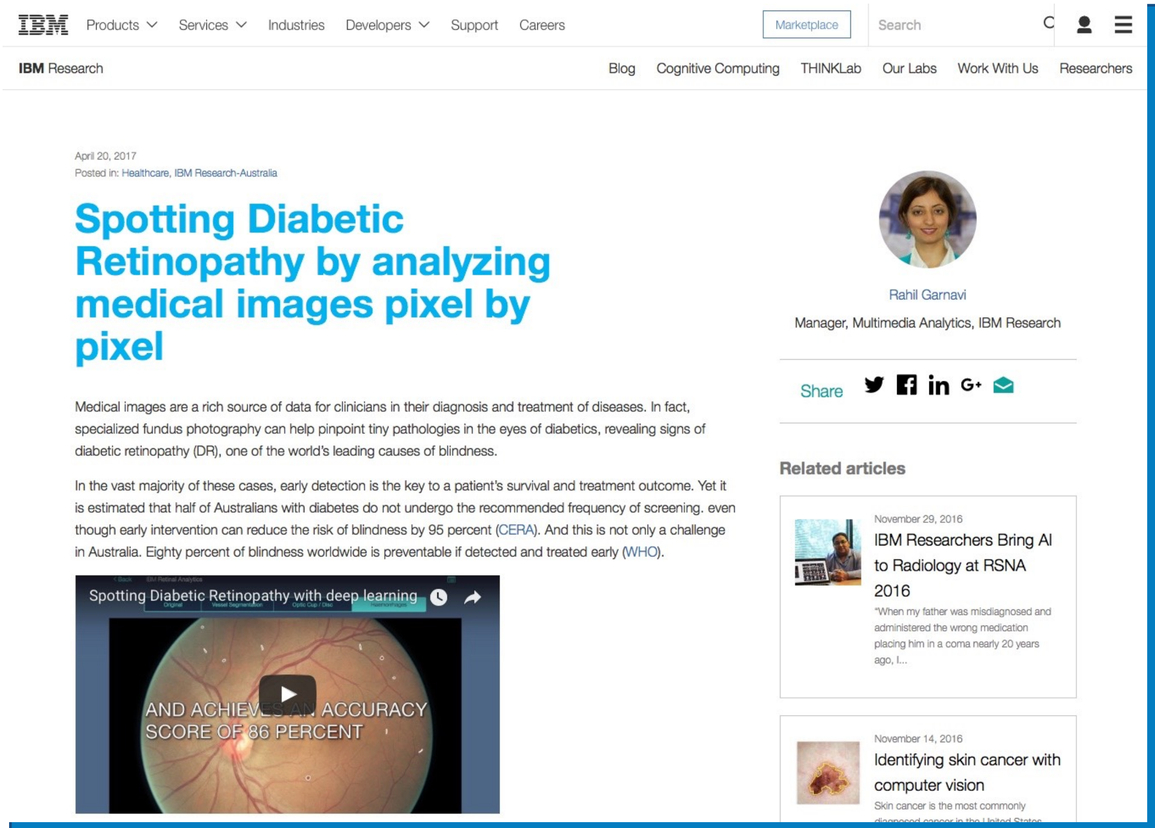
Posted December 19, 2016 By Firozgar Hansotia 0 70

Hyderabad, 19 December 2016: Continuing to enhance real life impact through technology, Microsoft India, in collaboration with L V Prasad Eye Institute, today launched Microsoft Intelligent Network for Eyecare (MINE). This is a mission-driven global consortium of like-minded commercial, research and academic institutions who have joined hands to apply artificial intelligence to help in the elimination of avoidable blindness and scale delivery of eyecare services worldwide.

The partner organizations of this consortium include Bascom Palmer – University of Miami, Flaum Eye Institute – University of Rochester (USA), Federal University of Sao Paulo (Brazil) and Brien Holden Vision Institute (Australia). The partner organizations will collaborate and collectively work on diverse datasets of patients across geographies to come up with machine learning predictive models for vision impairment and eye disease. This will include the rate of change of myopia in children, conditions that impact children's eyesight, predictive outcomes of refractive surgery, optimal surgery parameters as well as ways to personalize a surgery and maximize its probability of success. By studying this data and applying advanced analytics with Microsoft machine learning technology to derive insights, MINE will aim to drive strategies to prevent avoidable blindness and help increase efficiency in the delivery of eyecare worldwide. Currently 285 million people are visually impaired, of which 55 million reside in India. As the technology driver, Microsoft will deploy its leading cloud platform technology – Cortana Intelligence Suite, for advanced analytics and to build Artificial Intelligence models on eyecare.

STORM INSIGHTS INC.

Diabetic Retinopathy: IBM



"Our technology combines two various analytics approaches into one hybrid method, wherein a **convolutional neural networks (CNN)-based method for DR classification** is integrated with a **dictionary-based learning that incorporates DR-specific pathologies**. This hybrid analysis resulted in a great improvement in classification accuracy. Our method **takes approximately twenty seconds to analyze the image and achieves an accuracy score of 86 percent in classifying the disease across the five severity levels**.

An eye scan with Diabetic Retinopathy hemorrhages highlighted
We also showed a new method for the accurate segmentation of the fovea in retinal color fundus images. Fovea is responsible for sharp central vision. Location of retina lesions and pathologies with respect to the fovea impacts their clinical relevance. Our technology allows for a pixel-wise segmentation of the fovea. It **does not require prior knowledge of the location of other retinal structures such as optic disc or retina vasculature.** This is an advantage over other published methods, which can either localize only the center of the fovea or need a priori information about major structures in the image."

STORM INSIGHTS INC.

Transforming Healthcare & Pharma: Precision Medicine

Data-Driven
Personalized Treatment

Massachusetts Institute of Technology

HOME DAVID SONTAG PEOPLE RESEARCH PUBLICATIONS SOFTWARE TEACHING

Precision Medicine

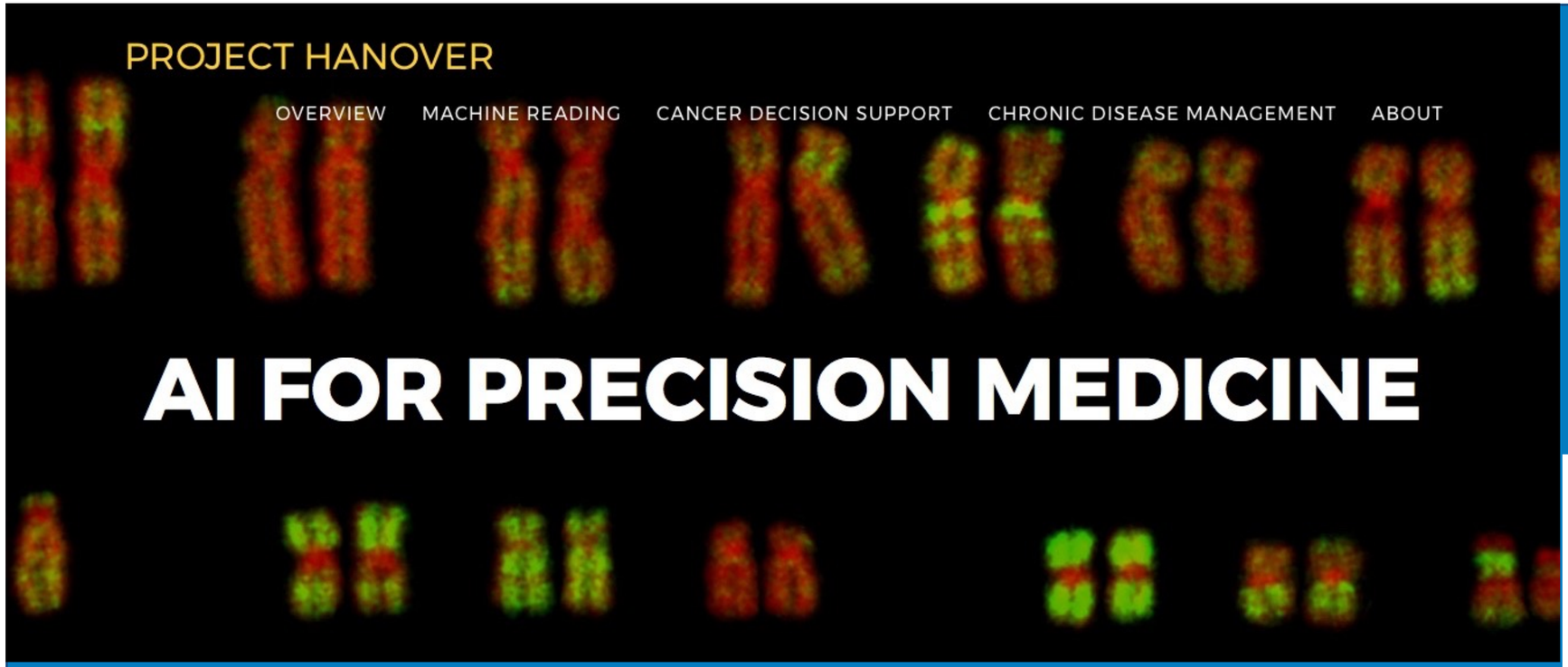
These are exciting times for the practice of medicine. The rapid adoption of electronic health records and has created a wealth of new data about patients, which is a goldmine for improving our understanding of human health. Our lab develops algorithms that use this data to better understand disease progression and to facilitate new, precise treatment strategies for a wide range of diseases and conditions such as Type 2 diabetes, which affects tens of millions of people worldwide every year, and multiple myeloma, a rare blood cancer. In pursuit of these aims, a major methodological focus has been on developing novel approaches to modeling high-dimensional time-series data, particularly approaches that bring together probabilistic modeling and deep learning, and causal inference from observational data.

- R. Krishnan, U. Shalit, D. Sontag. **Structured Inference Networks for Nonlinear State Space Models**, *Thirty-First AAAI Conference on Artificial Intelligence*, Feb. 2017. [code]
- N. Razavian, J. Marcus, D. Sontag. **Multi-task Prediction of Disease Onsets from Longitudinal Laboratory Tests**. Proceedings of the 1st *Conference on Machine Learning in Health Care* (MLHC), Aug. 2016. [code]
- X. Wang, D. Sontag, F. Wang. **Unsupervised Learning of Disease Progression Models**. *ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (KDD), Aug. 2014. [Slides]
- N. Razavian, S. Blecker, A.M. Schmidt, A. Smith-McLallen, S. Nigam, D. Sontag. **Population-Level Prediction of Type 2 Diabetes using Claims Data and Analysis of Risk Factors**. *Big Data* (Data and Healthcare Special Issue), Jan. 2016 [PDF]

http://clinicalml.org/research.html

STORM INSIGHTS INC.

Transforming Healthcare & Pharma: Microsoft



http://hanover.azurewebsites.net

We envision that AI-powered decision support for precision medicine will become an explosive growth area in cloud-based health analytics. In Project Hanover, building on prior work in Literome, we are making progress in three directions:

- **Machine reading**: We are developing NLP technology for converting text into structured databases, which enables us to build genome-scale knowledge bases by automatically reading millions of biomedical articles.
- **Cancer decision support**: We are collaborating with the Knight Cancer Institute to develop AI technology for cancer precision treatment, with current focus on developing a machine learning approach to personalize drug combinations for Acute Myeloid Leukemia (AML), where treatment hasn't improved in the past three decades.
- **Chronic disease management**: Our long-term goal is to develop AI technology for predictive and preventive personalized medicine to combat the soaring cost in caring for cancer and other chronic diseases, which accounts for nearly 90% of the U.S. healthcare spending.

STORM INSIGHTS INC.

TRANSFORMING HEALTHCARE: MICROSOFT

Microsoft Store Products Support

Official Microsoft Blog | The Fire Hose | Microsoft On the Issues | Next | Transform

Microsoft and partners combine the cloud, AI, research and industry expertise to focus on transforming health care

Posted February 16, 2017 By **Peter Lee** - *Corporate Vice President, Microsoft Research NExT*

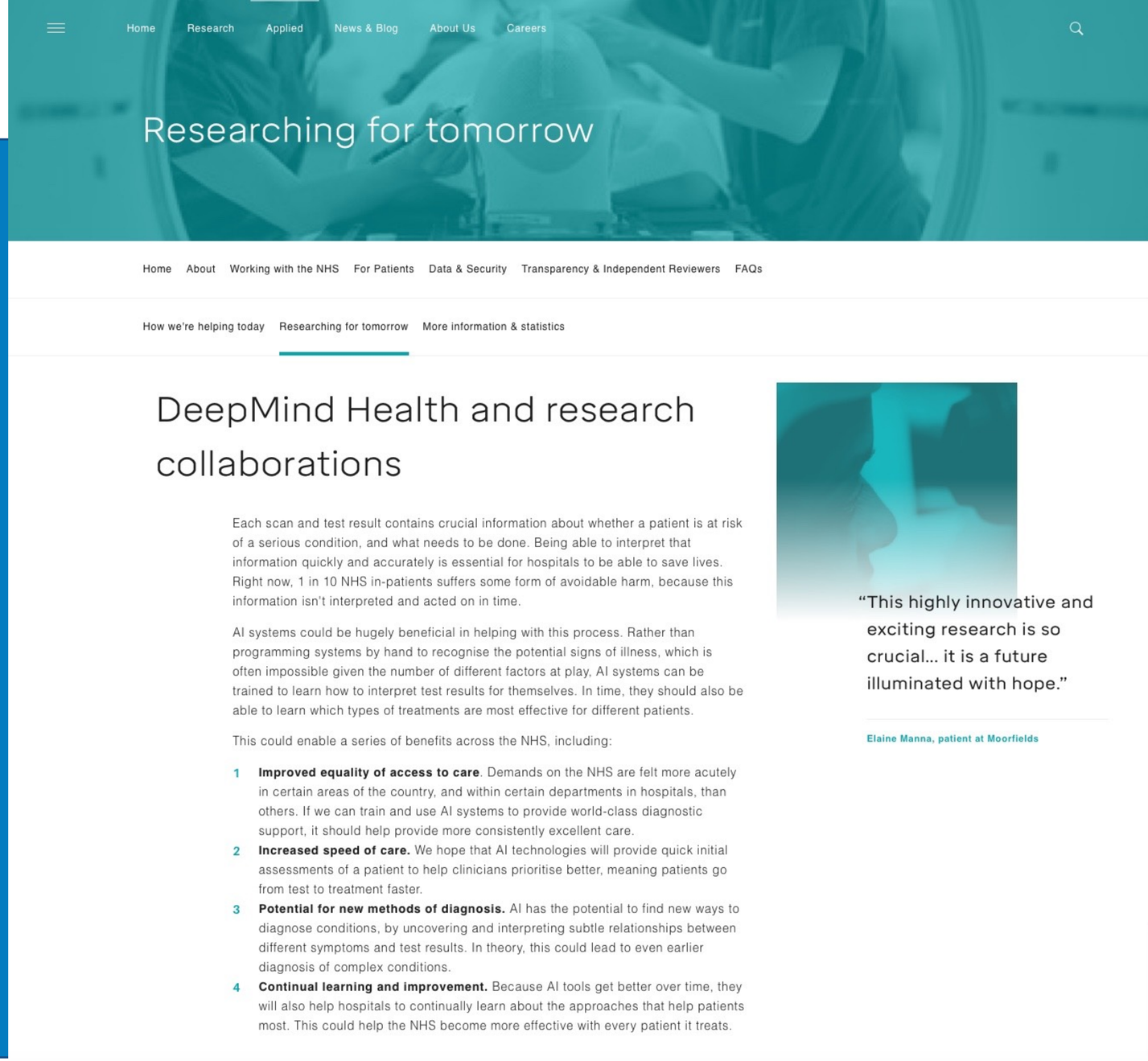
Microsoft announces new health-focused initiatives and solutions, and emerging strategic research partnership with UPMC

The goal is noble: Empower people to lead healthier lives. And yet, few industries in the world face more complex problems than health care. Disparate and disconnected information systems, the uncertainties within regulatory environments around the world and the inevitable disruptions in core business models all pose perplexing and interlocking challenges.

As we look at some of the challenges in health care, a natural question we ask ourselves is, How can Microsoft bring its capabilities to bear to solve some of these problems?

It's a big challenge. But we believe technology – specifically the cloud, AI and collaboration and business optimization tools – will be central to health care transformation.

TRANSFORMING HEALTHCARE: GOOGLE



STORM INSIGHTS INC.

Transforming Healthcare & Pharma: IBM

IBM Watson Health

Solving health challenges

IBM Watson for Genomics

Bringing the promise of precision medicine to more cancer patients, Watson can interpret genetic testing results faster and with greater accuracy than manual efforts. Our partnership with Quest Diagnostics means that all providers can potentially benefit, regardless of access to in-house sequencing.

Watson for Genomics

IBM Watson for Drug Discovery

Help researchers identify novel drug targets and new indications for existing drugs. The platform can help researchers uncover new connections and develop new treatments ahead of the competition.

Learn more about Drug Discovery

IBM Watson Health Patient Engagement

Identify patients with care gaps and automate personalized interventions, keeping patients engaged and helping them manage their own care between visits.

Learn more about Patient Engagement

IBM Watson for Oncology

Spend less time searching literature and more time caring for patients. Watson can provide clinicians with evidence-based treatment options based on expert training by Memorial Sloan Kettering (MSK) physicians.

Watson for Oncology

IBM Watson Care Manager

Use personalized care plans, automated care management workflows, and integrated patient engagement capabilities to help create more informed action plans.

Learn more

STORM INSIGHTS INC.

TRANSFORMING HEALTHCARE: IBM

Memorial Sloan Kettering Cancer Center

Giving | Our Locations | Find a Doctor

For Adult Patients | For Child & Teen Patients | For Healthcare Professionals | For Research Scientists

IBM Watson and Quest Diagnostics Launch Genomic Sequencing Service Using Data from MSK

On October 18, IBM Watson Health and Quest Diagnostics announced the launch of IBM Watson Genomics from Quest Diagnostics, a new service that helps advance precision medicine by combining cognitive computing with genomic tumor sequencing. Memorial Sloan Kettering will supplement Watson's corpus of scientific data with **OncoKB**, a precision oncology knowledge base, to help inform individual treatment options for cancer patients.

The new service involves laboratory sequencing and analysis of a tumor's genomic makeup to help reveal mutations that can be associated with targeted therapies and clinical trials. Watson then compares those mutations against relevant medical literature, clinical studies, pharmacopeia, and carefully annotated rules created by leading oncologists, including those from MSK. Watson for Genomics ingests approximately 10,000 scientific articles and 100 new clinical trials every month.

Bolstering the corpus of data Watson uses, MSK will provide OncoKB, a database of clinical evidence that will help Watson uncover treatment options that could target the specific genetic abnormalities that are causing the growth of the cancer. Comparison of literature that may take medical experts weeks to prepare can now be completed in significantly less time.

OncoKB was developed and is maintained through MSK's Marie Josée and Henry R. Kravis Center for Molecular Oncology, in partnership with Quest. It includes annotation for almost 3,000 unique variants in 418 cancer-associated genes and in 40 different tumor types, including descriptions of the effects of specific mutations as well as therapeutic implications.

The project is publicly accessible, meaning that researchers around the world have access to information about oncogenic effects and treatment implications of thousands of unique variants at their fingertips.

"We now know that genetic alterations are responsible for many cancers, but it remains challenging for most clinicians to deliver on the promise of precision medicine since it requires specialized expertise and a time-consuming interpretation of massive amounts of data," said Paul Sabbatini, Deputy Physician-in-Chief for Clinical Research at MSK. "Through this collaboration, oncologists will have access to MSK's expertly curated information about the effects and treatment implications of specific cancer gene alterations. This has the power to scale expertise and help improve patient care."

STORM INSIGHTS INC.

Transforming the Insurance Industry

Integrate/Combine Data Sources

Historical

Streaming - Telematics, GPS, Biometrics

Orthogonal - Weather, News, Social

New ML-enabled Business Models

Price by usage and behavior

Perhaps give the user feedback for 60-90 days before the data is shared with the insurer

AUTO: The weather on your route is stormy, your policy will be adjusted accordingly. Proceed?

LIFE: If that ice cream is for you, your life insurance policy will be adjusted after the glucose monitor report. Would you like to see an estimate of the real cost of the ice cream?

P&C: If Proposition 21 passes at the polls next week, your risk of flooding will rise and your policy will be adjusted accordingly.

STORM INSIGHTS INC.

Keep in Touch

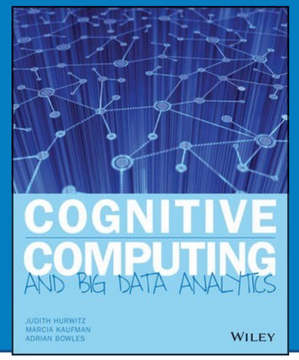
adrian@storminsights.com
adrian@aragonresearch.com

Twitter @ajbowles
Skype ajbowles

If you would like to connect on LinkedIn, please let me know that you that you registered for the Smart Data webinar series.

Upcoming SmartData Webinar Dates & Topics

- ☑ July 13 Advances in NLP I: Understanding
- ☑ August 10 Organizing Data and Knowledge: The Role of Taxonomies and Ontologies
- ☑ Sept. 14 Advances in Natural Language Processing II: NL Generation



New Content from Aragon Research

AragonResearch.com

STORM INSIGHTS INC.

DATAVERSITY®

Aragon Research

Resources for Further Research 1

- http://blogs.royalsociety.org/in-verba/2016/10/05/machine-learning-in-the-pharmaceutical-industry/
- http://www.huffingtonpost.com/adi-gaskell/using-machine-learning-to_b_12049046.html
- https://www.techemergence.com/applications-machine-learning-in-pharma-medicine/
- https://www.technologyreview.com/s/604271/deep-learning-is-a-black-box-but-health-care-wont-mind/?utm_content=buffera2da5&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer

STORM INSIGHTS INC.

Resources for Further Research 2

- https://deepmind.com/applied/deepmind-health/working-nhs/health-research-tomorrow/ Clinical Trials
- https://cloud.google.com/blog/big-data/2017/03/using-machine-learning-for-insurance-pricing-optimization
- http://news.mit.edu/2016/faster-gene-expression-profiling-drug-discovery-0128
- https://research.googleblog.com/2015/03/large-scale-machine-learning-for-drug.html
- http://thevarsity.ca/2017/02/27/accelerating-drug-discovery-with-machine-learning/
- http://news.stanford.edu/2017/04/03/deep-learning-algorithm-aid-drug-development/
- https://www.ibm.com/blogs/research/2017/04/spotting-diabetic-retinopathy/
- http://ijarcet.org/wp-content/uploads/IJARCET-VOL-4-ISSUE-12-4415-4419.pdf
- http://ijarcet.org/wp-content/uploads/IJARCET-VOL-4-ISSUE-12-4415-4419.pdf

STORM INSIGHTS INC.