
Machine learning techniques for predicting complications and evaluating drugs efficacy

Three examples of real world evidence studies: combating AIDS, predicting diabetes complications and the epilepsy trial and error challenge

Dr. Michal Rosen-Zvi
Senior manager
Analytics department
IBM Research lab - Haifa, Israel

Outline

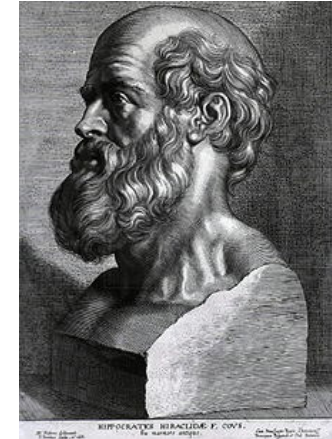
- Data and needs in the medical world
- Machine Learning in a nutshell
- Three concrete examples
 - combating AIDS
 - predicting diabetes complications
 - the epilepsy trial and error challenge
- Wrap up

The Medical domain

The first known medical record was developed by Hippocrates, in the fifth century B.C. He prescribed two goals:

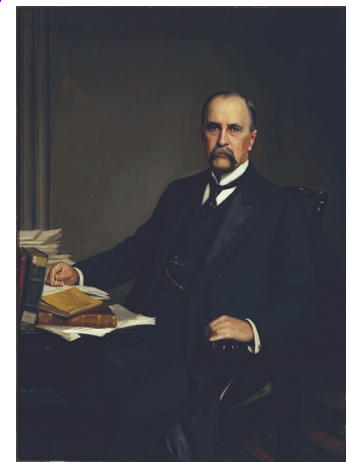
- A medical record should accurately reflect the course of disease
- A medical record should indicate the probable cause of disease

Hippocrates of Cos (460 BC – 370 BC)
Father of western medicine

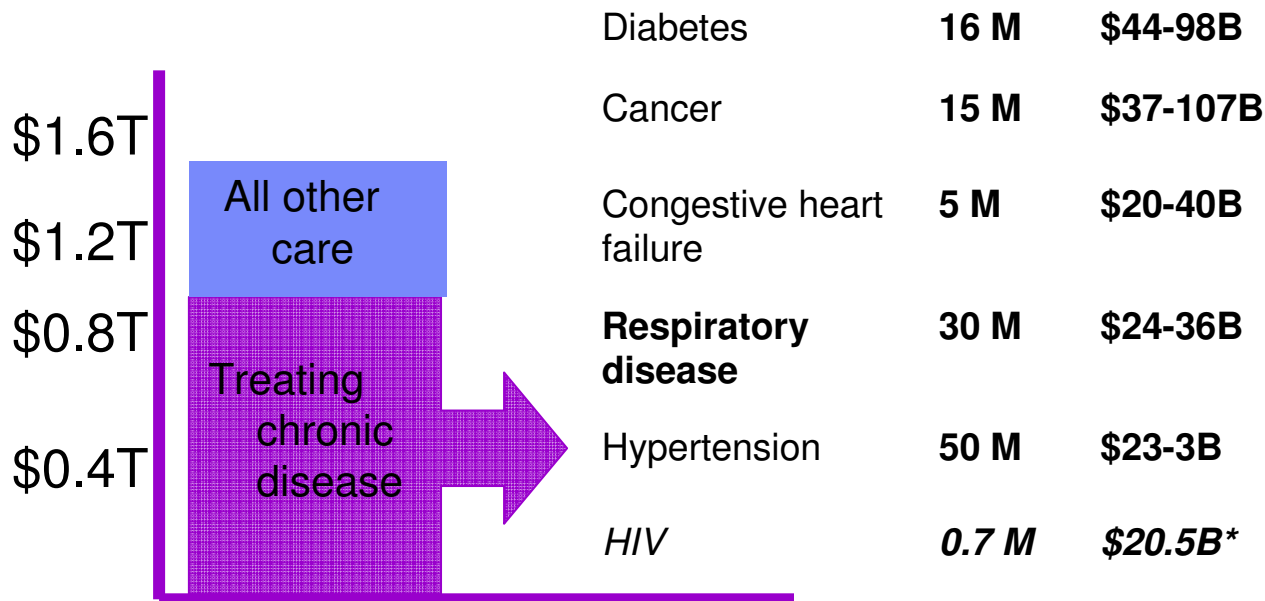


“Medicine is a science of uncertainty and an art of probability.”
“The greater the ignorance the greater the dogmatism.
The truth of one age become the absurdities of the next, and the foolishness of yesterday become the wisdom of tomorrow.”

Sir William Osler, (1849 – 1919),
Father of modern medicine



Chronic diseases – prevalence and cost



US healthcare spending 2002

Personalized healthcare requires Machine Learning capabilities for improving the care in various ways, including:

- **Matching between experts to patients (specialist referral)**
- **Prediction of spending per person per current status of the patient**
- **Assisting physician to optimize treatment based consumable analytics**
- **Assisting pharmacy companies in clinical trials recruitments, results analysis**

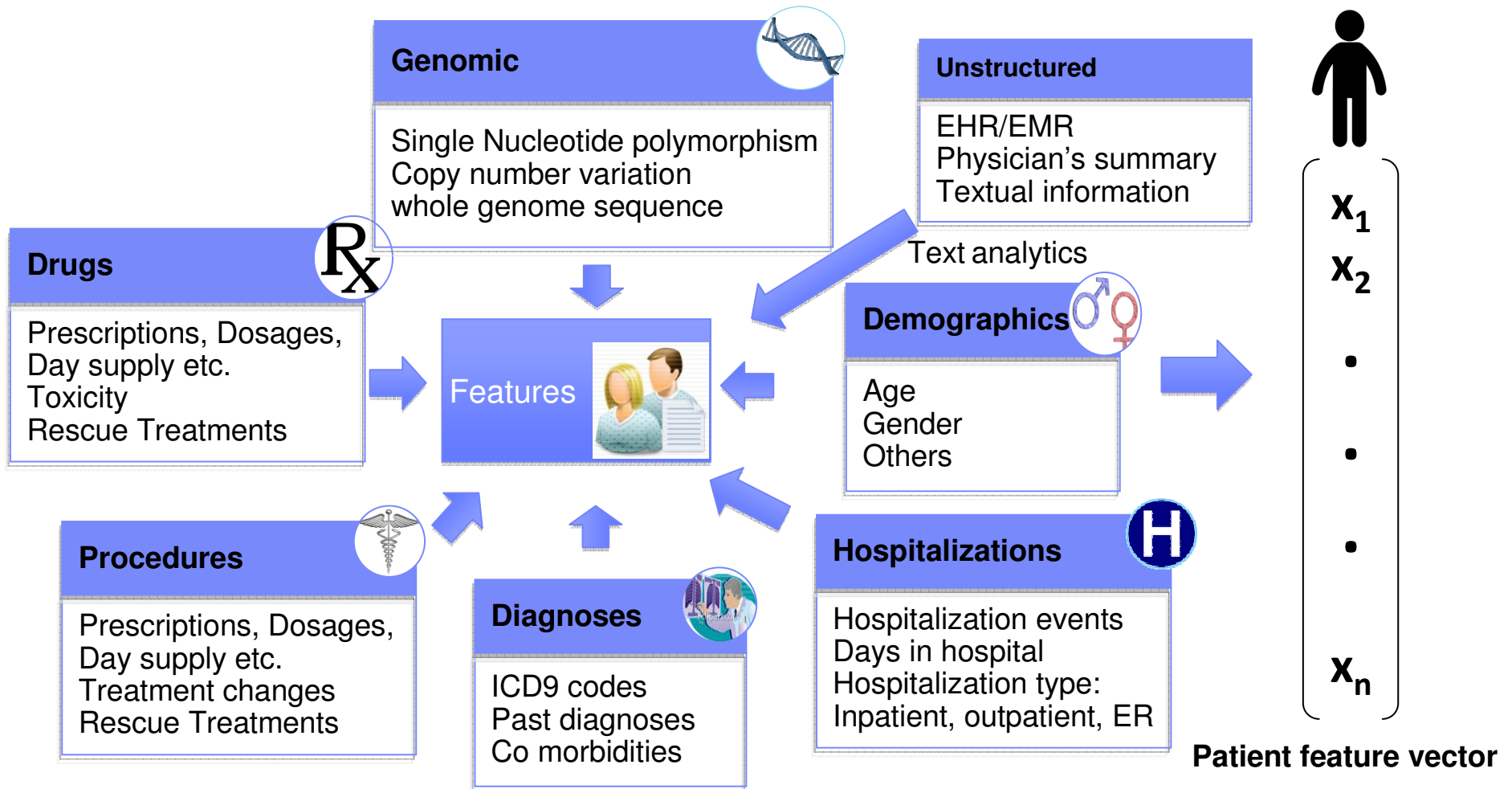
What is real world evidence?

- The availability of huge data sets from the point of care has led to real-world evidence (RWE) for healthcare
- Pharmaceutical companies can use RWE to
 - Deepen the understanding of diseases and treatments
 - Better direct the design of products
 - Improve mechanisms to recruit patients for clinical-trials
 - Gain and defend market access
 - Lead innovative smart solutions such as
 - Clinical decision support for personalized treatments
 - Care coordination for treating mentally ill patients, who are prone to attrition
 - Personalized dosage optimization systems



Prediction models are based on the data features

- A patient is represented as a vector of features



RWE Potential value per sector

- Research and development (R&D)
 - Modeling is a key feature for a more integrated approach that can accelerate novel drug development
 - Analytics of large cohorts of data might lead to findings of off label usage and effectiveness of drugs and get fast-tracked into clinical trials in alternative indications
 - Phase I: Analytics of what characterizes patients for who existing drugs are not effective might lead to novel findings of novel drugs
 - Phase III: Deep analytics of existing drugs and risk factors of patients might assist in cohort selection and in designing treatments, in particular when a cocktail of existing drugs and experimental drugs is considered

- marketing activities
 - Analytics might lead to automatic identification of characteristics of patients for whom drugs are most effective and enable more targeted marketing
 - Adherence problems might be inferred from the data and addressed by marketing activities



Watson for Epilepsy UCB-IBM announcement





16 May, 2013


European
pharmaceutical
review

PharmaTimes ONLINE

UCB and IBM bet on big data to help epilepsy

WORLD NEWS | MAY 17, 2013


SHARE    ...  +1 0

 Tweet 1

KEVIN GROGAN

UCB has teamed up with IBM to "harness super-computer power" and help improve treatment for epilepsy patients.

The partners have announced the completion of the initial phase of a "pioneering project designed to harness the power of super computers and intelligent analytics" aimed at delivering personalised care for people with epilepsy. They add that "this exciting milestone marks the critical first step in the path towards eventually harnessing the power of cognitive computing capabilities", for patients.



Could a 'super computer' help improve epilepsy care? Yes, it's elementary my dear WATSON

Publication date: 16 May 2013

Author: UCB

Tagged with: Iris Löw-Friedrich, UCB

UCB and IBM have today announced the completion of the initial phase of a pioneering project designed to harness the power of super computers and intelligent analytics with the aim of helping healthcare providers deliver personalized care for people with epilepsy. This exciting milestone marks the critical first step in the path towards eventually harnessing the power of cognitive computing capabilities, such as IBM's Watson, for patient care.

The two companies are working together to create the most comprehensive bank of scientific literature and patient data on epilepsy ever collated. Both organisations hope this initiative will enable healthcare professionals from around the world to access the latest information and use a more informed approach to patient care. Upon completion of this project, healthcare professionals will be able to use their own clinical patient assessment with the system's predictive analytics to deliver more personalized and specific approaches to care will be successful.

The Street

UCB And IBM Collaborate To Personalize Care For Epilepsy Patients

BY [PRNewswire](#) | 05/16/13 - 09:47 AM EDT

Stock quotes in this article: [IBM](#)

Find out if (IBM) is in Cramer's Portfolio. ▶



 PRNewswire

ARMONK, N.Y. and BRUSSELS, May 16, 2013 /PRNewswire/ -- **regulated information** -- UCB and IBM (NYSE:IBM) today announced the completion of the initial phase of a project designed to harness the power of analytics to help healthcare providers deliver more highly personalized care to people living with epilepsy. The milestone marks the critical first step in the path towards eventually harnessing the transformative power of cognitive computing capabilities, such as IBM Watson, for epilepsy care.

Outline

- Data and needs in the medical world
- Machine Learning in a nutshell
- Three concrete examples
 - combating AIDS
 - predicting diabetes complications
 - the epilepsy trial and error challenge
- Wrap up

Machine Learning

- Learning from data samples

Statistics;
Data Mining

Supervised Learning

- Samples are labeled

Unsupervised;
Semi-supervised

Classification

- The labels represent association with one of a few classes

Regression;
Ranking

Passive Learning

- The learner cannot select samples to label

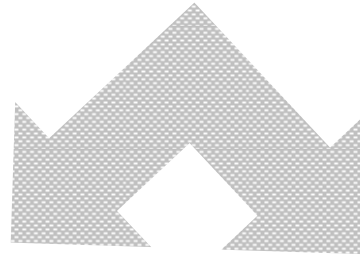
Active
Learning

Batch Learning

- Training is performed independently of the testing

Online
Learning

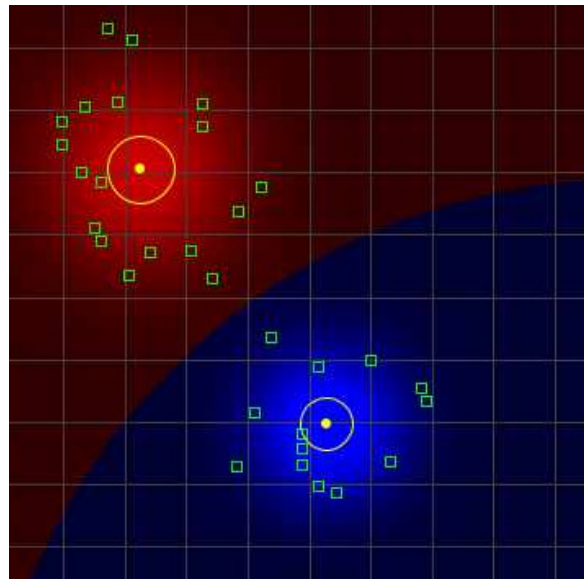
Classification Algorithms



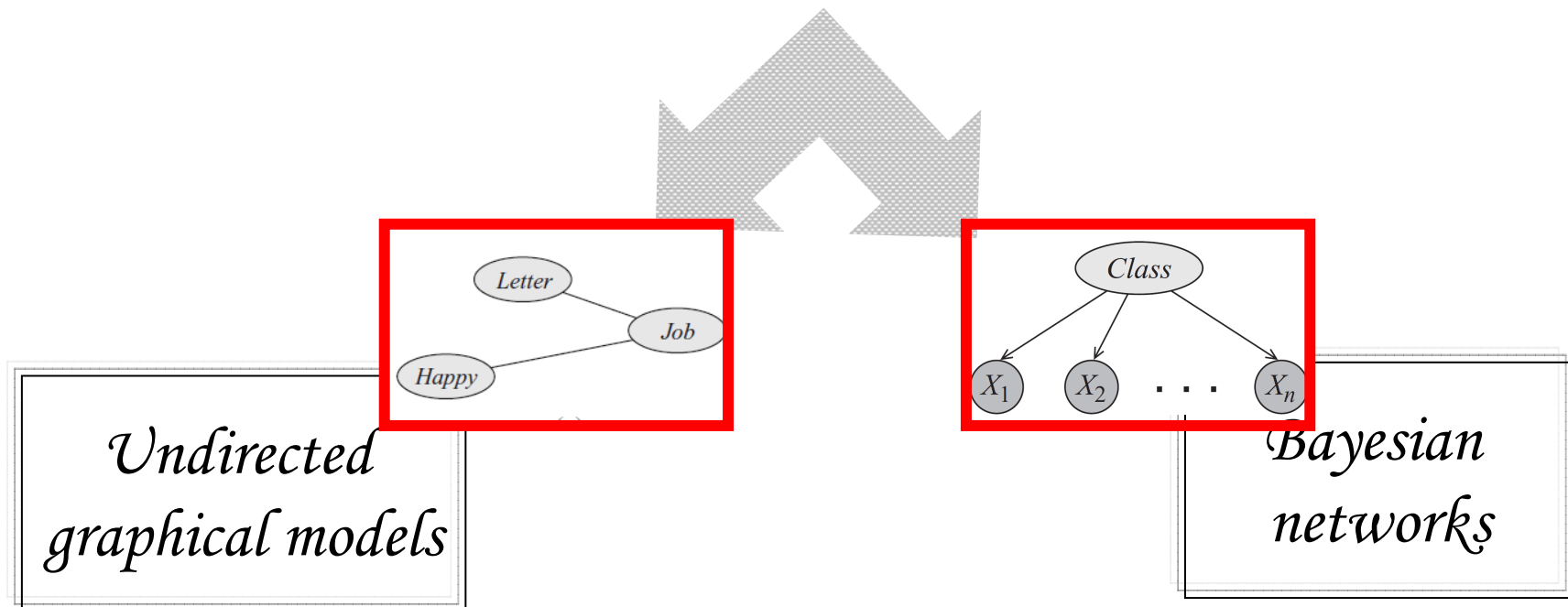
*Discriminative
Methods*



*Generative
Methods*



Probabilistic graphical models

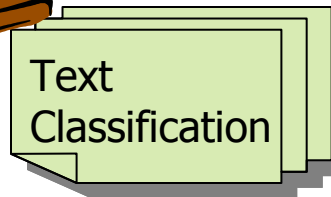
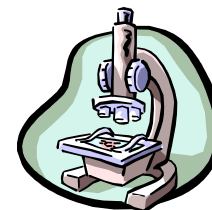
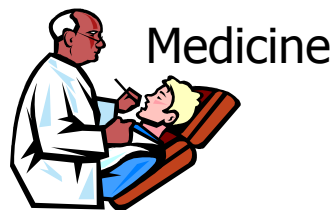
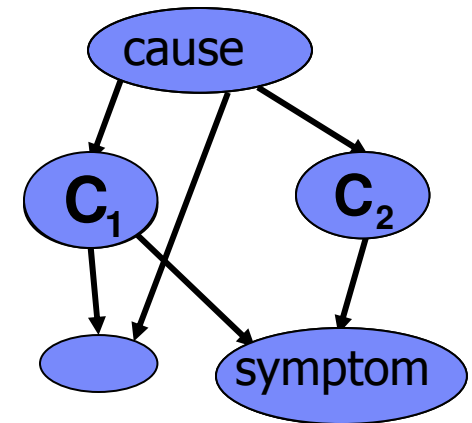


Probabilistic graphical models use a graph based representation as the basis for compactly encoding a complex distribution over a high dimensional space

Three aspects: Representation, learning, inference

What are Bayesian Networks useful for?

- Diagnosis: $P(\text{cause}|\text{symptom})=?$
- Prediction: $P(\text{symptom}|\text{cause})=?$
- Classification: $\max_{\text{class}} P(\text{class}|\text{data})$
- Decision-making (given a cost function)



Rules of Probability

- Sum rule and product rule

$$P(A) = \sum_B P(A, B)$$

$$P(A, B) = P(A|B)P(B)$$

- Bayes rule

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

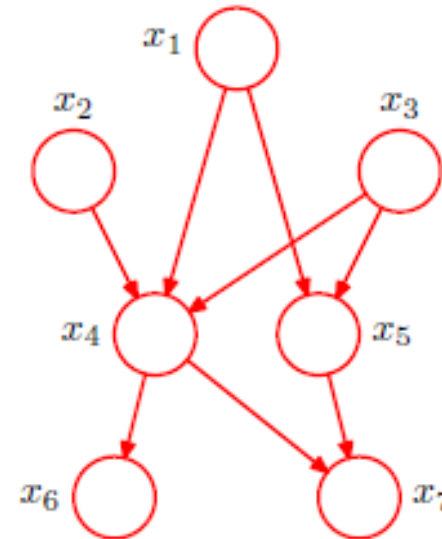
Bayesian Networks

- The term "Bayesian networks" was coined by Judea Pearl in 1985 and formally introduced in "Probabilistic reasoning in Intelligence Systems" (1988)
- A Bayesian network is defined using a directed graph and the properties of the directed graph imply the conditional independence properties of the BN
- Definitions:
 - A directed graph is a pair (V,E) where V is a set of nodes and E a set of ordered pairs of vertices or directed edges
 - A directed acyclic graph is a directed graph that has no directed edges
 - Given three random variables X, Y, Z we say that X is conditionally independent of Y given Z if for all x, y, z values it holds that
$$p(x,y|z)=p(x|z)p(y|z)$$
 - The parents of i are all the nodes with edges going into i denoted as $Pa(i)$
 - A Bayesian network is a model of a distribution that is constructed via a DAG

Factorization properties

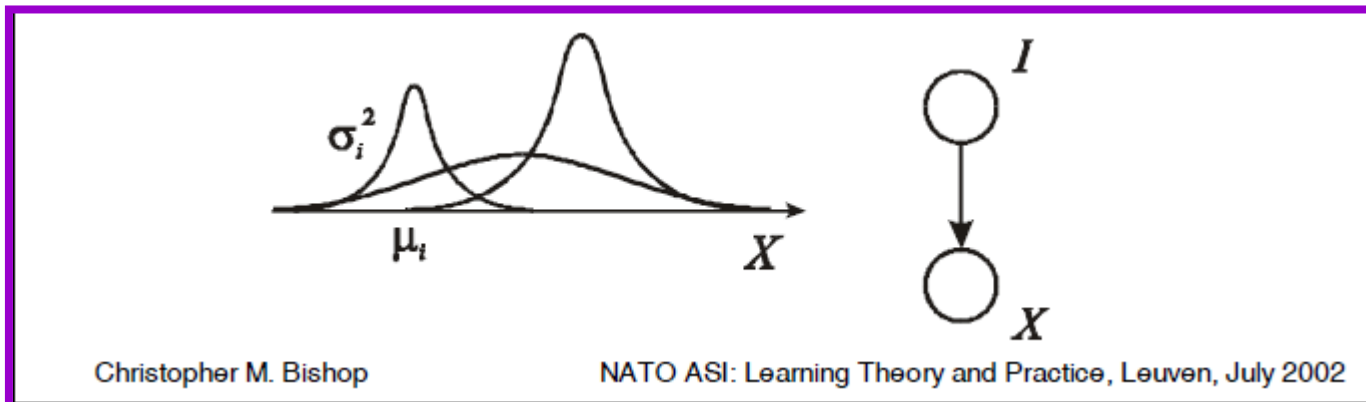
- DAGs : there exists an ordering of the nodes such that there are no links that go from any node to any lower numbered node

$$P(x) = \prod_{i=1}^K p(x_i | pa(i))$$



Example: Mixture of Gaussians

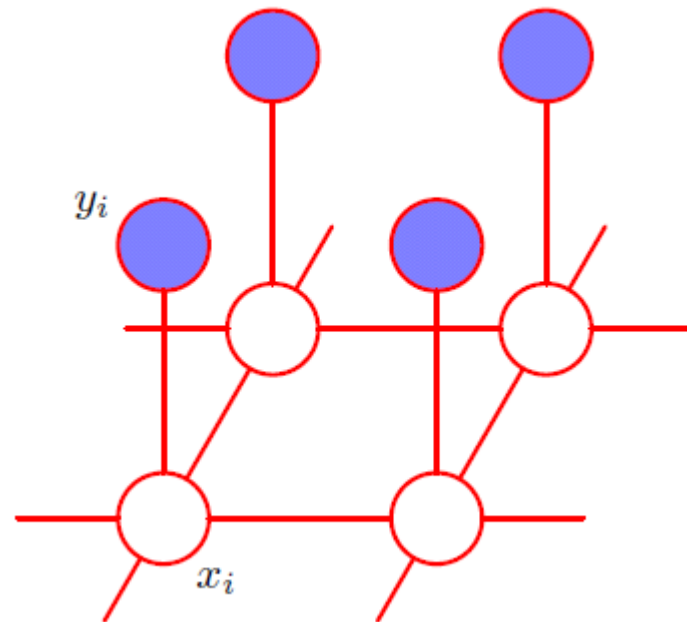
- $P(\text{Class}=i)=\pi_i$
- $P(X|\text{class}=i)=\text{Normal}(\mu_i, \sigma_i)$
- $P(X)=\sum_i P(\text{Class}=i) P(X|\text{class}=i) = \sum_i \pi_i \text{Normal}(\mu_i, \sigma_i)$



Undirected graphical models

- Markov random field is defined local interactions between random variables and is provided by a product of factors
- Image de-noising is one of the first examples of using Markov Random Field
- Besag, 1974; Geman and Geman, 1984; Besag, 1986

$$P_{\varphi}(X_1, X_2, \dots, X_b) = \frac{1}{Z} \phi_1(D_1) \phi_2(D_2) \dots \phi_m(D_m)$$



Outline

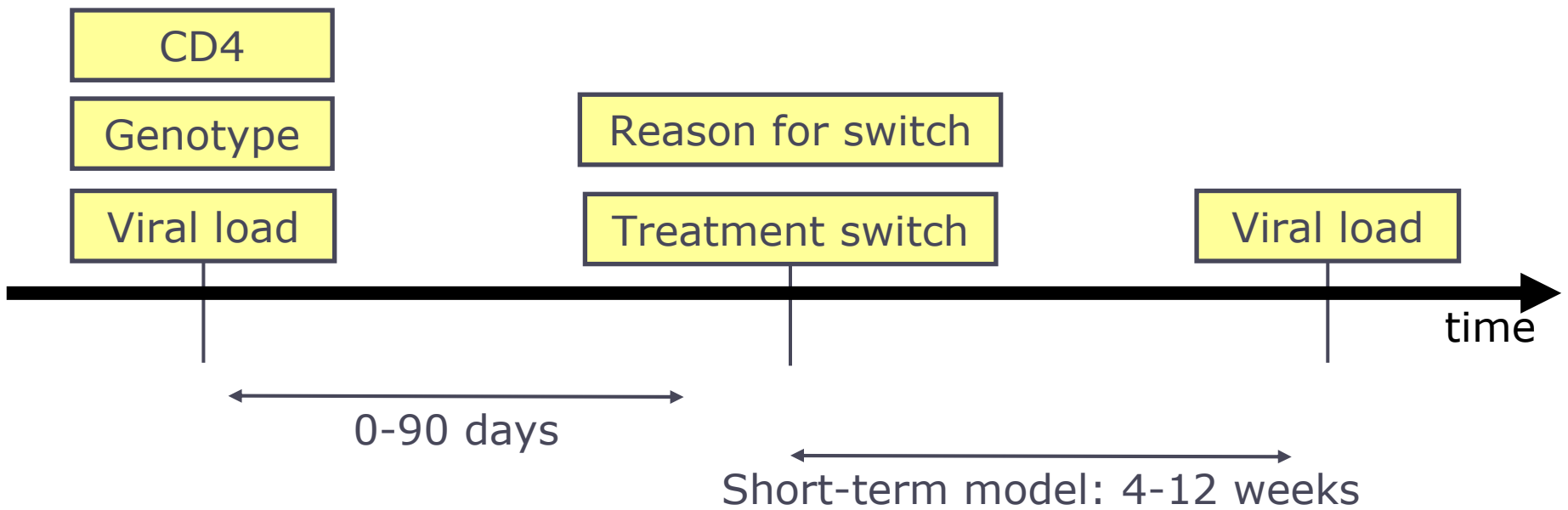
- Data and needs in the medical world
- Machine Learning in a nutshell
- Three concrete examples
 - combating AIDS
 - predicting diabetes complications
 - the epilepsy trial and error challenge
- Wrap up

EuResist Partners: a combined effort of Virologists, IT and ML experts

- Informa S.r.l., Italy, Rome
- Università degli Studi di Siena, Italy, Siena
- Karolinska Institutet, Sweden, Stockholm
- University of Cologne – University hospital, Germany, Cologne
- Max-Planck-Society for Advancement of Science, Germany Saarbrücken
- IBM Haifa Research Lab
- Research Institute for Particle and Nuclear Physics of the Hungarian Academy of Sciences, Hungary, Budapest
- European Federation of Pharmaceutical Industries and Associations (EFPIA)
- Kingston University, United Kingdom, Kingston upon Thames
- University of Roma TRE, Italy



From EMRs and blood tests to standard datum for automatic learning



- Patient demographics (age, gender, race, route of infection)
- Past AIDS diagnosis
- Past treatments
- Past genotypes

Predictive analytics – Therapy optimization

EuResist: Predict efficacy of anti-retroviral drug regimens against given HIV

- World's largest database centered on HIV resistance and clinical response information
- Prediction models analyze information from:
 - 60,000 patients
 - 150,000 therapies
 - 500,000 viral load measurements
- Prediction engines are 76% accurate
- Outperforms alternative automatic-prediction engines and team of human experts
- Nine out of ten experts made more mistakes than our prediction engines

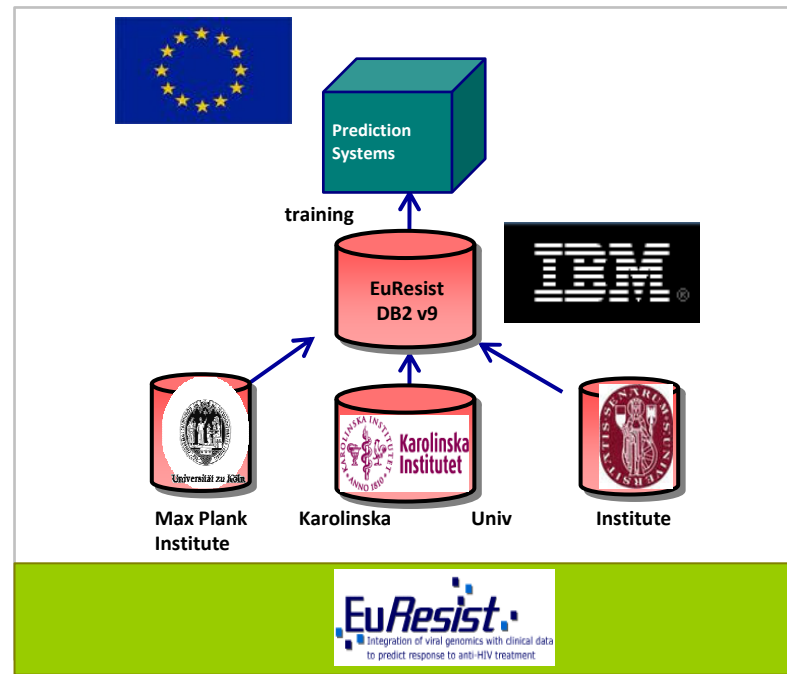
Labeling therapies and prediction engines

- **A therapy needs to at least contain a baseline viral load at therapy switch and a follow-up measure**
- Label a therapy as a success or failure based on follow-up viral load, learn from the past to predict on a new one how likely that is a success
- **Three engines are trained and predict outcome**
 - The **Evolutionary Engine** uses mutagenetic trees to compute the genetic barrier to drug resistance
 - The **Generative Discriminative Engine** employs a Bayesian network modeling interactions between current and past antiretroviral drugs
 - The **Mixture of Effects Engine** includes second and third-order variable interactions between drugs and mutations

Predictive analytics – Therapy optimization

EuResist: Predict efficacy of anti-retroviral drug regimens against given HIV

- Data and deep analytics to improve HIV/AIDS therapy
- Based on viral genotype data integrated with treatment response
- IBM hosts and manages the largest HIV clinical genomics database in the world
- Freely available on-line clinical support system since 2008
- Contract with GEIE – an NOP European Economic Interest Group
- Recent studies on adherence inference



- **Value**

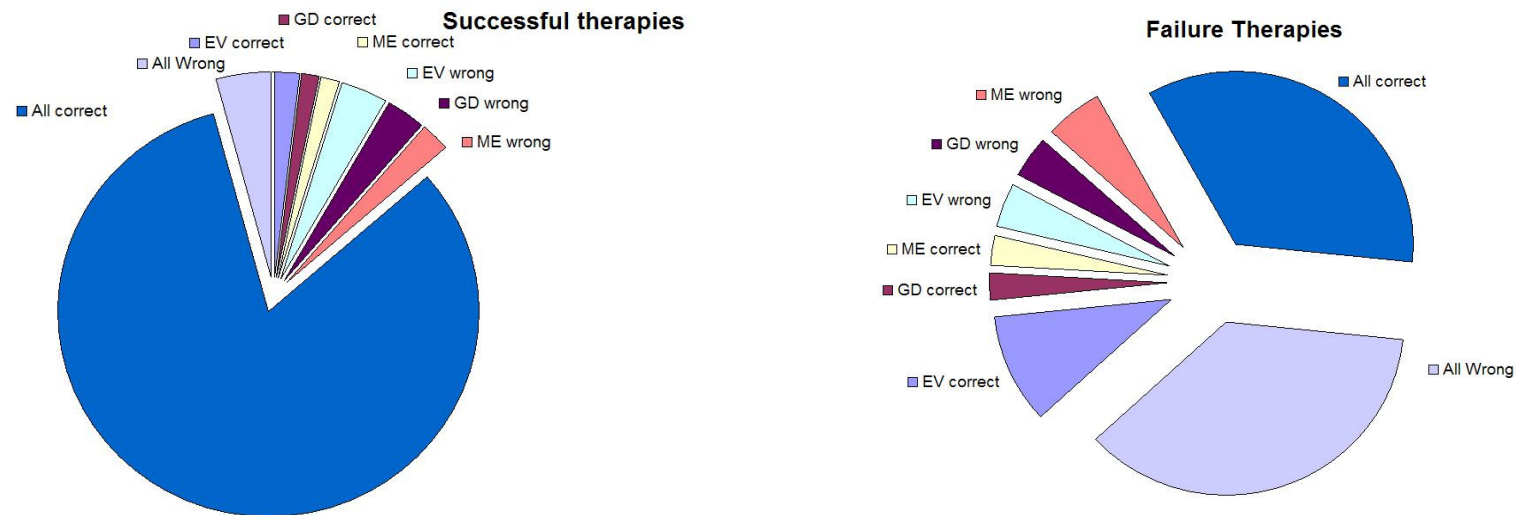
- Patients cohort analysis
- More sophisticated patient recruiting methods
- Data analysis can assist in the detection of patients with adherence issues

- **Selected References**

- J. Bogojeska et al. History-alignment models for bias-aware prediction of virological response to HIV combination therapy. *Journal of Machine Learning Research - Proceedings Track 22*: 118-126 (2012)
- M. Zazzi et al. Prediction of response to antiretroviral therapy by human experts and by the EuResist data-driven expert system (the EVE study). *HIV Medicine*, 2010
- Prosperi et al. (2010) Antiretroviral Therapy Optimisation without Genotype Resistance Testing: A Perspective on Treatment History Based Models. *PLoS ONE* 5(10)
- Rosen-Zvi et al. (2008), "Selecting anti-HIV therapies based on a variety of genomic and clinical factors", *ISMB conference/ bioinformatics journal*

Comparison of performances

A comparison of the three engines prediction on failure or success therapy – where they fail or succeed together and where there is a single winner



“Happy families are all alike; every unhappy family is unhappy in its own way”

Leo Tolstoy, *Anna Karenina*, Chapter 1, first line

Outline

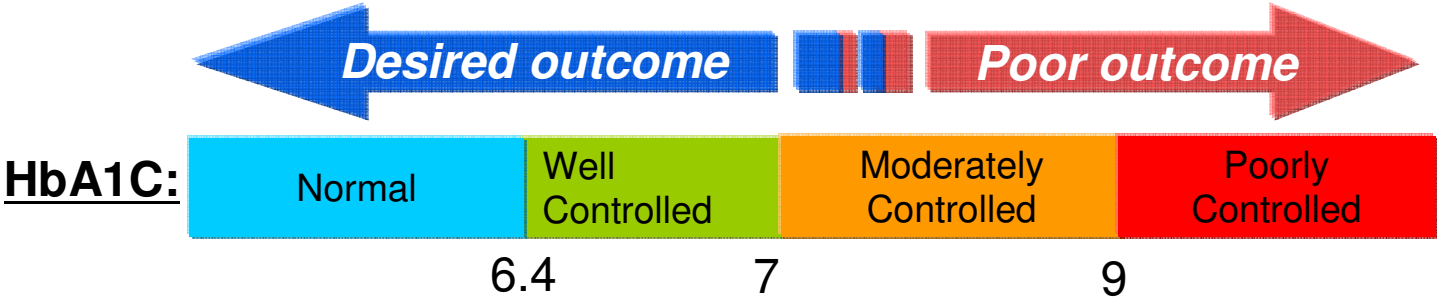
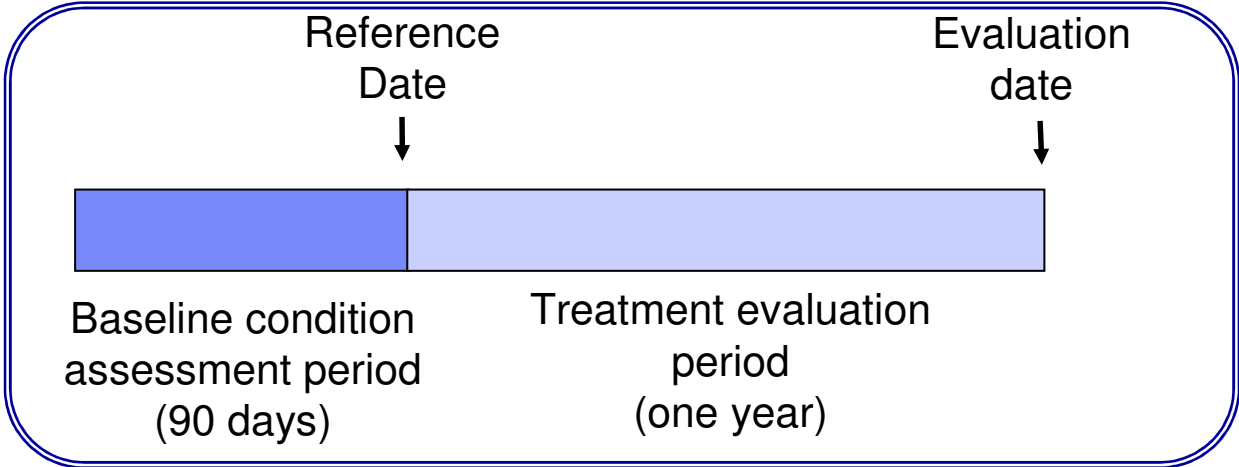
- Data and needs in the medical world
- Machine Learning in a nutshell
- Three concrete examples
 - combating AIDS
 - predicting diabetes complications
 - the epilepsy trial and error challenge
- Wrap up

Personalized Management of Chronic Diseases

- Diabetes:
 - Abundant in 7.8% of the population in the US, and rising
 - The leading cause of blindness, kidney failure, and non-traumatic lower limb amputations among adults.
- **Questions:**
 - *Which patients can gain from a personalized approach?*
 - *Can their treating physician significantly affect their condition?*



Data Representation



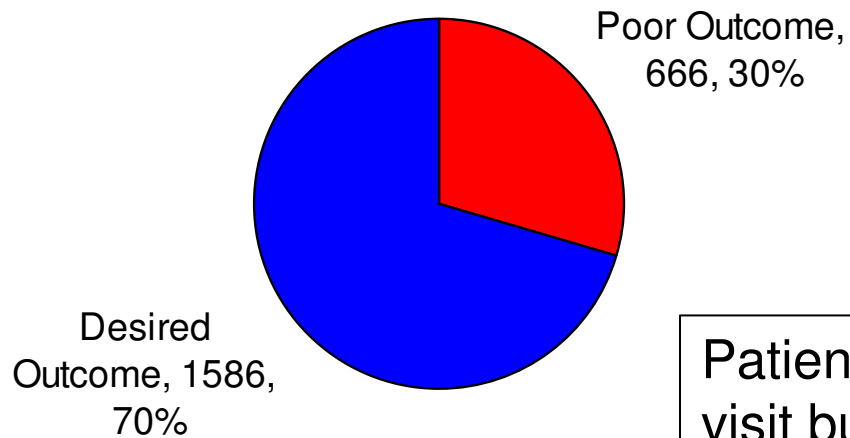
Data



HbA1C outcome prediction
2152 patients

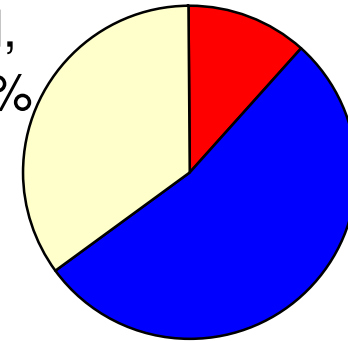


Urgent care outcome prediction
4545 patients



Censored,
1604, 35%

Patients with no UC
visit but less than 1Y
follow-up period

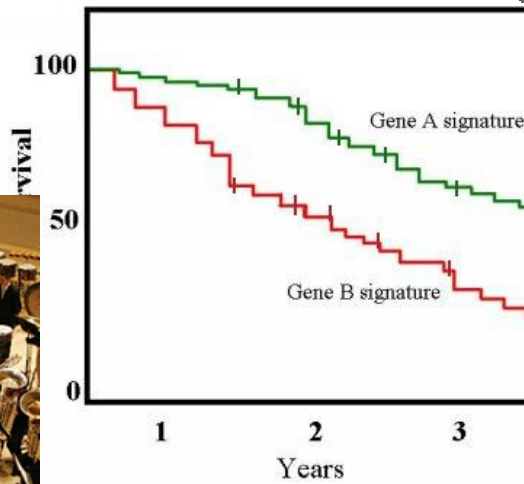
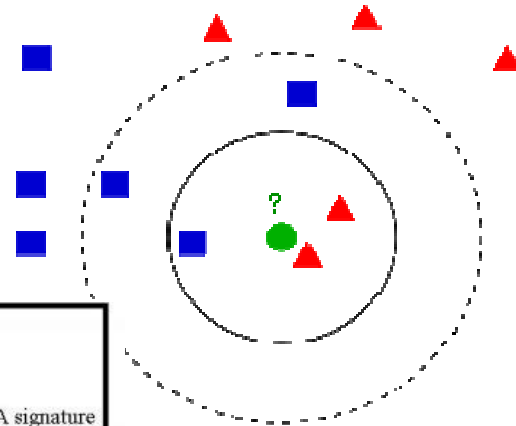
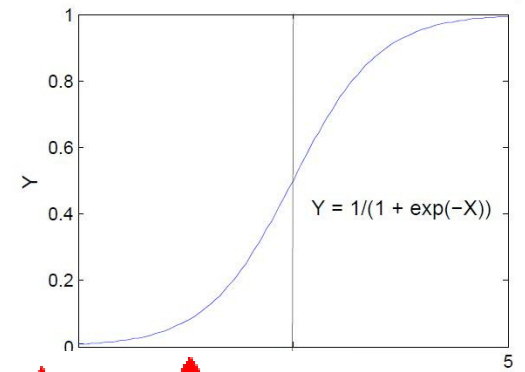


UC in 1Y,
541, 12%

No UC
Records,
2400, 53%

Learning Models

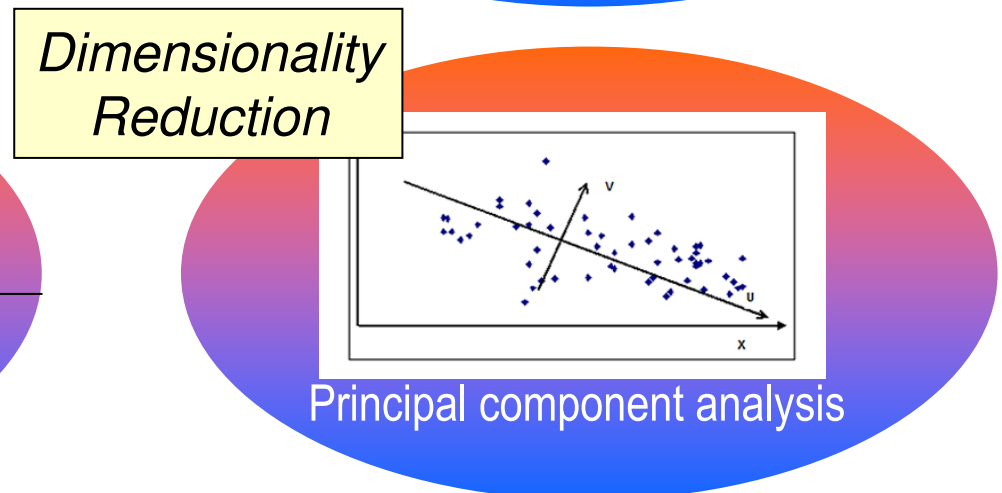
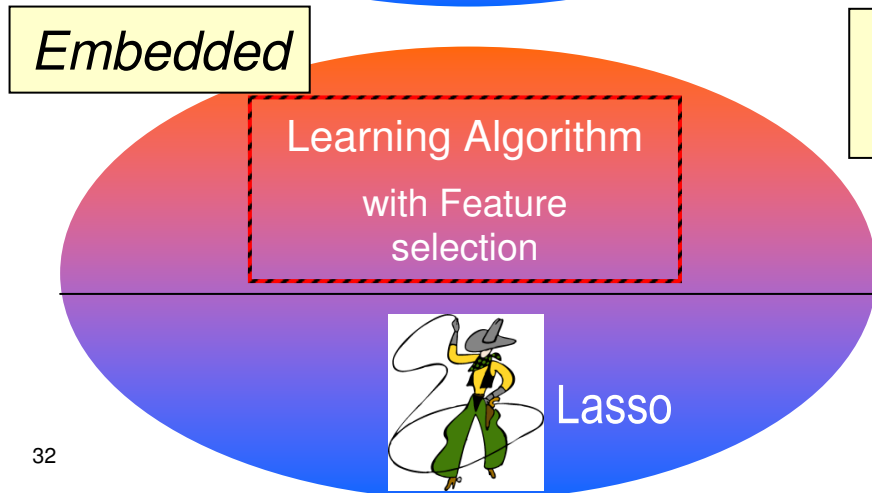
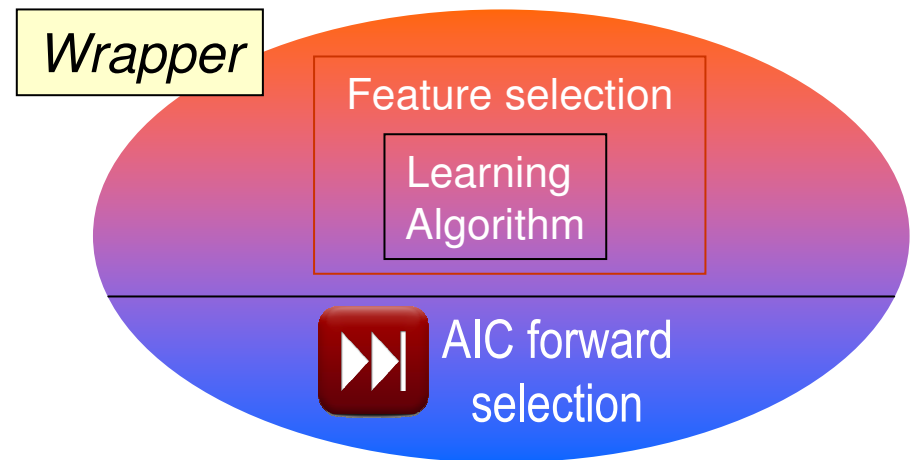
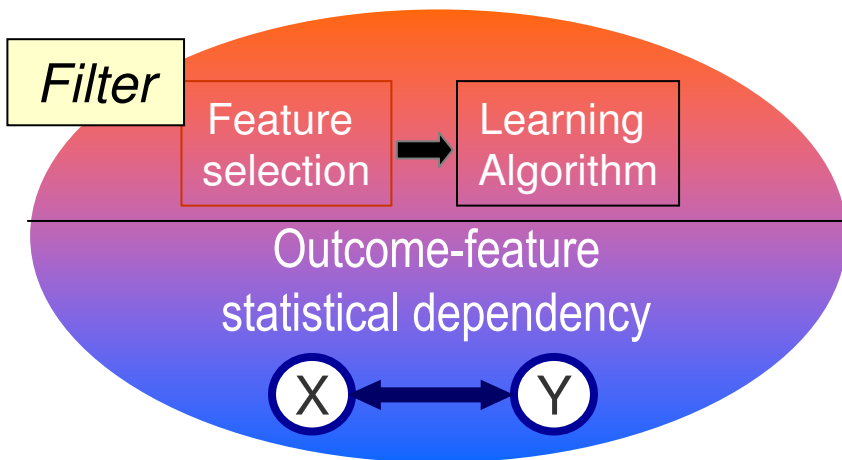
- Logistic regression
- Weighted k-nearest neighbors
- Cox proportional hazards
- Ensemble of models



Feature Selection and Dimensionality Reduction

▪ 2150 – 4500 patients, ~1000 features

→ dimensionality reduction is necessary to prevent over-fitting



Do Physicians Affect Patients' Outcome?

- Enhance with physician information:
 - Physician Features
 - Average treatment of the well/poorly managed population of a physician
 - Patient-2-Physician Similarity
 - Similarity of the patient to the well/poorly managed population of a physician at the baseline period

Poorly Managed



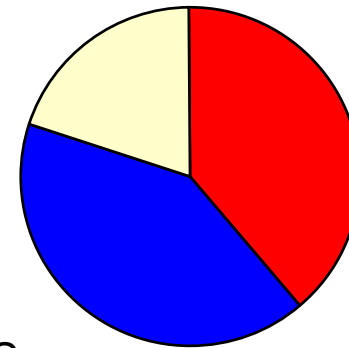
Well Managed



Patient-Physician Match – UC Visits

Focusing on the high-risk patients, physician features improve the prediction

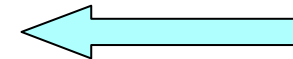
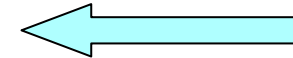
Censored, 40,
20%



UC in 1Y, 78,
39%

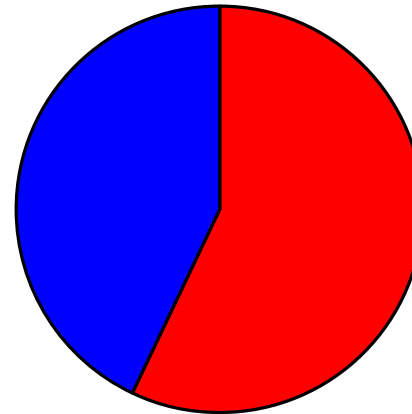
No UC
Records, 82,
41%

Feature Sets Used			C-Index		AUC	
Pat.	Phys.	P2P	mean	std	mean	std
0	1	0	0.51	0.06	0.52	0.11
1	1	0	0.54	0.05	0.48	0.10
1	0	0	0.56	0.04	0.48	0.05
0	0	1	0.60	0.02	0.53	0.05
0	1	1	0.60	0.02	0.53	0.05
1	1	1	0.61	0.03	0.53	0.05
1	0	1	0.61	0.02	0.54	0.05



Patient-Physician Match – HbA1C

Desired
Outcome, 43,
43%



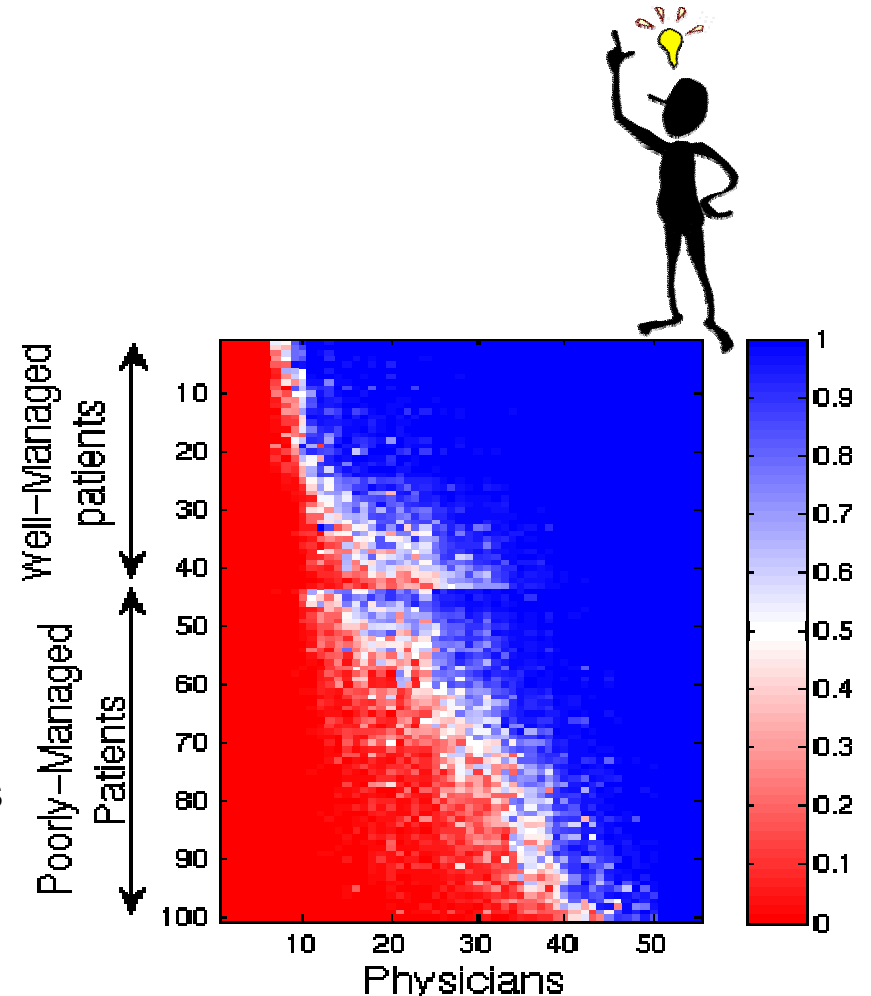
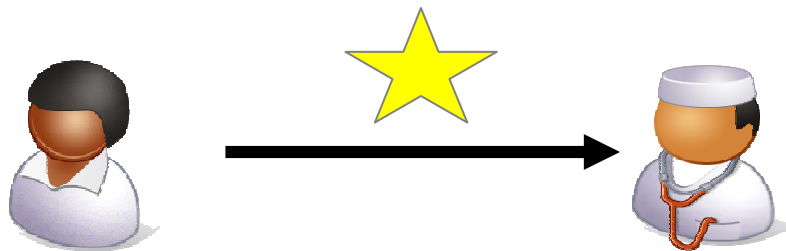
Poor
Outcome, 57,
57%

Feature Sets Used			AUC	
Pat.	Phys.	P2P	mean	std
<i>1</i>	<i>0</i>	<i>0</i>	<i>0.45</i>	<i>0.02</i>
0	1	0	0.49	0.06
1	1	0	0.49	0.06
1	0	1	0.56	0.03
0	0	1	0.56	0.03
1	1	1	0.59	0.03
<i>0</i>	<i>1</i>	<i>1</i>	<i>0.59</i>	<i>0.03</i>

Focusing on the high-risk patients physician features improve the prediction

What it is good for?

- Predict which (patient, physician) pairs has the best probability for treatment success!
 - Reveal best practice patterns of high-performing physicians
 - By investigating patient-physician features used by the model
 - Improve evaluation of physician work
 - Taking into account the case difficulty
 - Optimize the match between high-risk patients and physicians
 - Physician referral application



Predictions of treatment success for high-risk patients based on the HbA1c outcome.

Outline

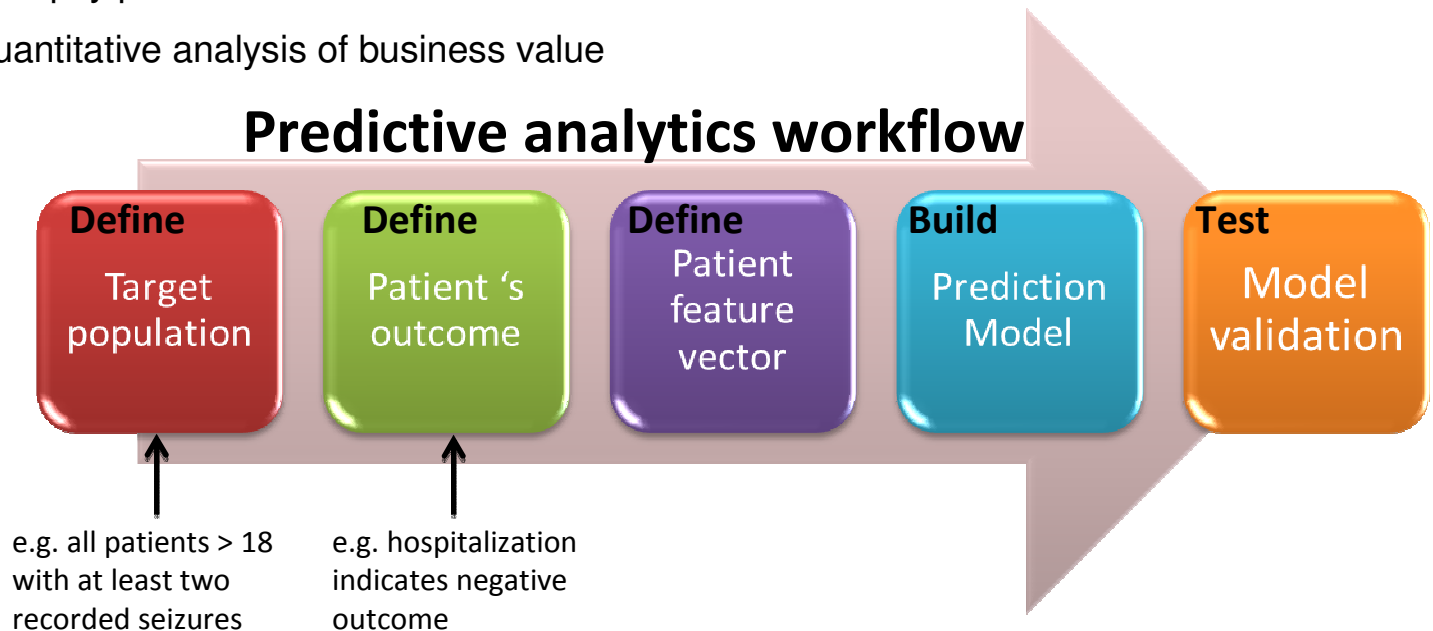
- Data and needs in the medical world
- Machine Learning in a nutshell
- Three concrete examples
 - combating AIDS
 - predicting diabetes complications
 - the epilepsy trial and error challenge
- Wrap up

Watson for Epilepsy

Pilot phase – patient similarity and predictive analytics in the Epilepsy domain

IBM deployed a complete workflow from data analysis to business value

- Starting from data definitions through exact and measurable outcome definition for Epilepsy patient
- Features extractions and expansion based on data as well as published literature and expert knowledge
- Patient similarity analysis
- Predictive analytics for building a prediction engine for predicting the best combination of drugs to provide to a particular epilepsy patient and that would lead to a better outcome
- Model validation and quantitative analysis of business value



Work performed w.

Hani Neuvirth, Michal Ozery-Flato, Yaara Goldschmidt, Ehud Aharoni, Ranit Aharonov, IBM Research – Haifa

Jianying Hu, Martin S. Kohn, Shahram Ebadollahi, IBM Watson Research lab

Jonathan Laserson Google



Contact information

Michal Rosen-Zvi,
rosen@il.ibm.com