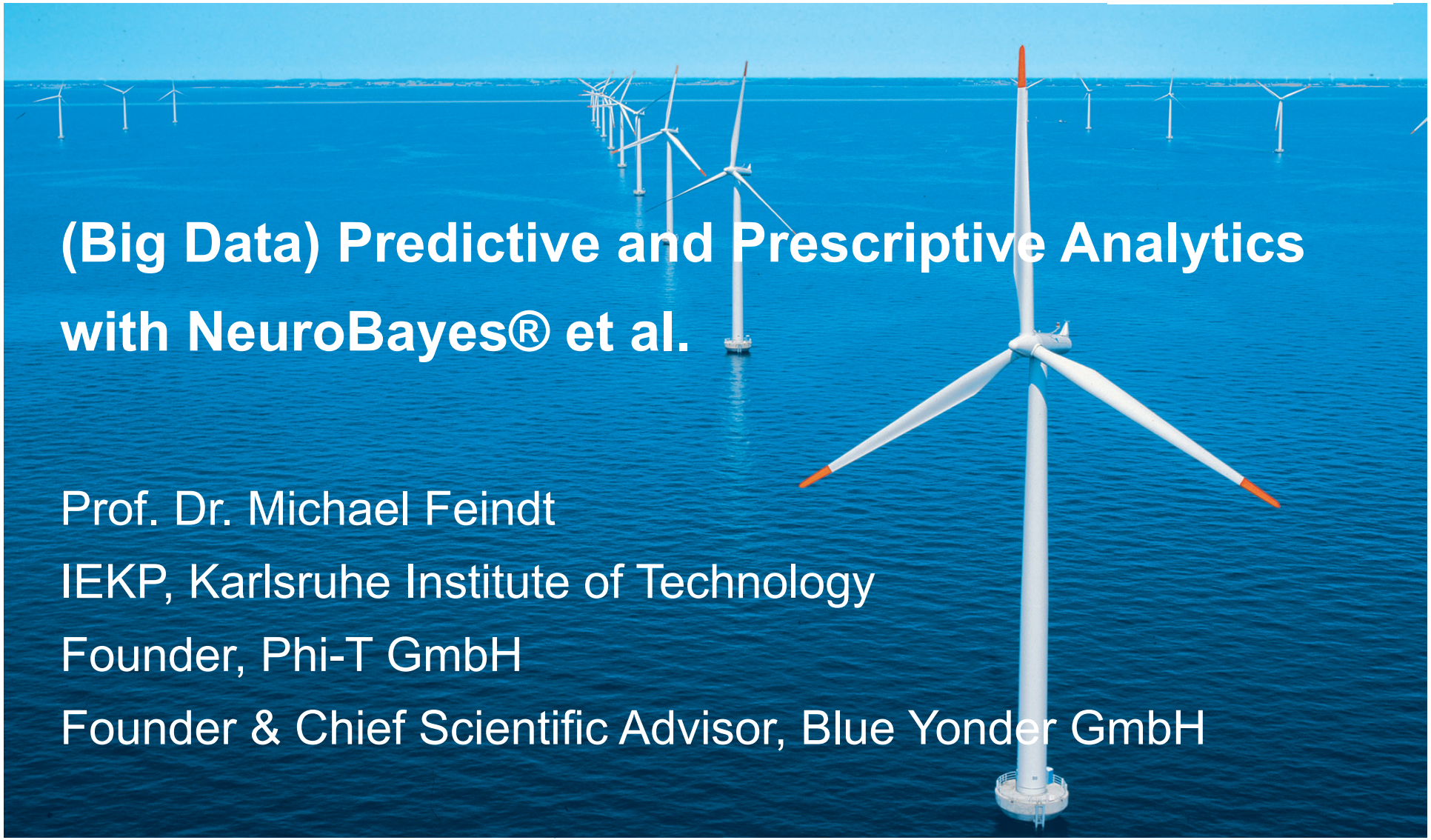
blue yonder
Forward looking. Forward thinking.

<phi-t>®
Physics Information Technologies

KIT
Karlsruhe Institute of Technology



(Big Data) Predictive and Prescriptive Analytics with NeuroBayes® et al.

Prof. Dr. Michael Feindt
IEKP, Karlsruhe Institute of Technology
Founder, Phi-T GmbH
Founder & Chief Scientific Advisor, Blue Yonder GmbH

blue yonder
Forward looking. Forward thinking.

Personal Vita

1982-1991 DESY (PLUTO, CELLO at PETRA)

1991-1997 CERN (DELPHI at LEP)

since 1997 Professor at Univ. Karlsruhe (now KIT)

since 1997 CDF II at Fermilab / DELPHI at LEP / (CMS at LHC)

since 2008 Belle, Belle II at KEK

1999/2000 invention of NeuroBayes algorithm

2002 foundation of Phi-T

2008 foundation of Blue Yonder,

with offices in Karlsruhe, Hamburg, London

Personal Vita

<phi-t>®
Physics Information Technologies

blue yonder
Forward looking. Forward thinking.

NeuroBayes task 1:
Classifications

Classification:

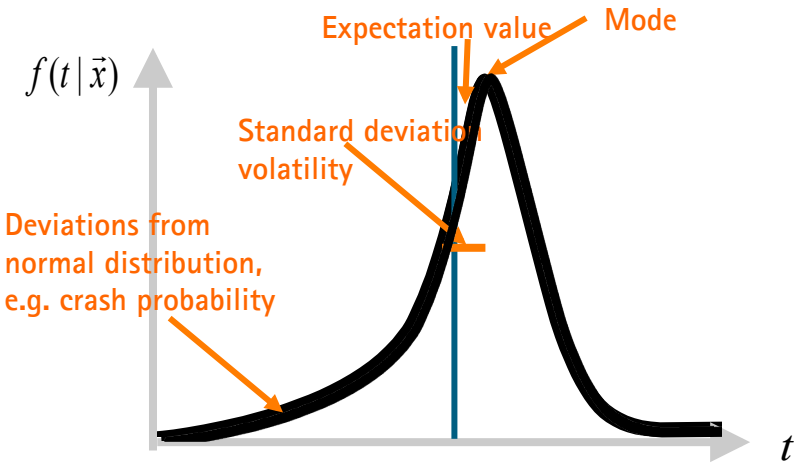
Binary targets: Each single outcome will be "yes" or "no"
NeuroBayes output is the Bayesian posterior probability that answer is "yes" (given that inclusive rates are the same in training and test sample, otherwise simple transformation necessary).

Examples:

> This elementary particle is a K meson.
> This jet is a b-jet.
> This three-particle combination is a D+.
> This event is real data and not Monte Carlo.
> This neutral B-meson was a particle and not an antiparticle at production time.
> Customer Meier will cancel his contract next year.

NeuroBayes task 2:
Conditional probability densities

Probability density for real valued targets:
For each possible (real) value a probability (density) is given. From that all statistical quantities like mean value, median, mode, standard deviation, etc. can be deduced.



Examples:

> Energy of an elementary particle
(e.g a semileptonically decaying B meson with missing neutrino)
> Q value (invariant mass) of a decay
> Lifetime of a decay
> Phi-direction of an inclusively reconstructed B-meson in a jet.
> Turnaround of an article next year

(very important in industrial applications)

Roots: 30 years of elementary particle physics, peaking at the LHC at CERN.

Built to understand how exactly our universe works.

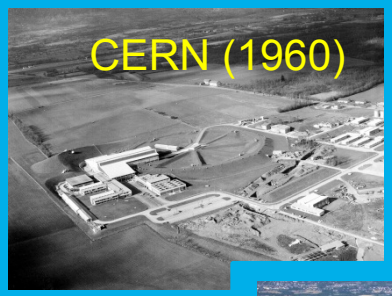


Photo: CERN

Background: High Energy Physics

Fundamental research at the forefront of science

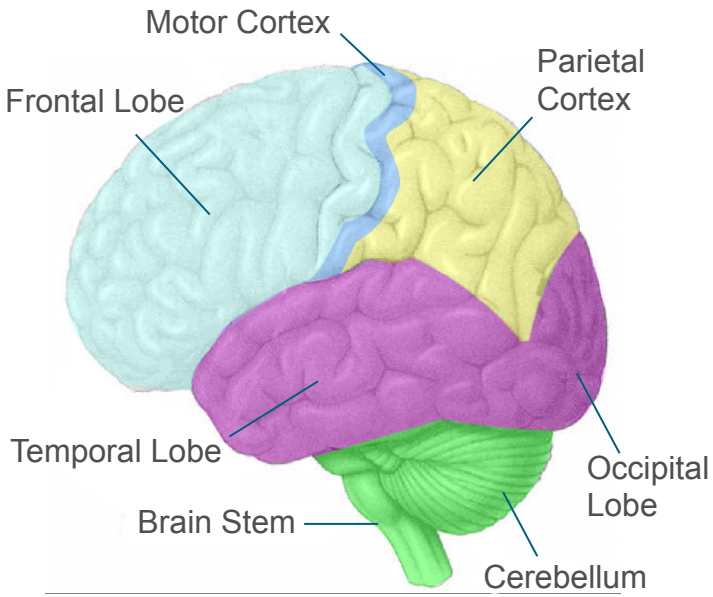
Very strong worldwide competition on getting results from data
→very strong statistical methods: fast & robust multivariate algos

Photo: CERN

One way to construct a one dimensional test statistic from multidimensional input (a MVA-method):

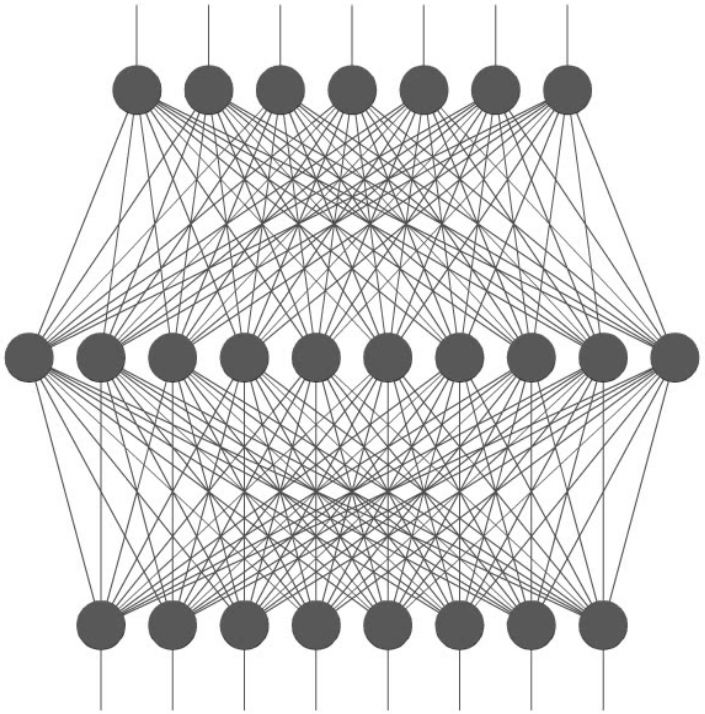
Neural networks

Self learning procedures, copied from nature



Neural networks

The NeuroBayes classification core is based on a simple feed forward neural network.
The information (the knowledge, the expertise) is coded in the connections between the neurons.



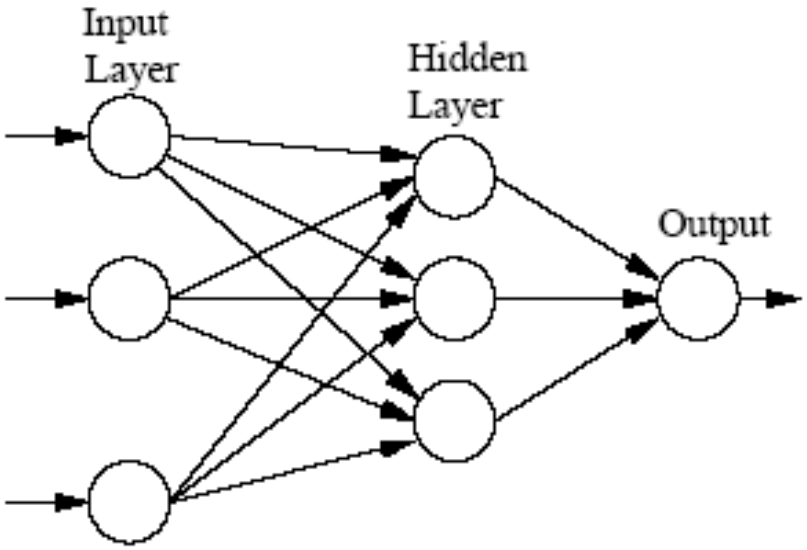
Each neuron performs fuzzy decisions.

A neural network can learn from examples. Supervised machine learning.

Human brain: about 100 billion (10^{11}) neurons
about 100 trillion (10^{14}) connections
NeuroBayes : 10 to few 100 neurons

Neural Network

basic functions



The output of node j in layer n is calculated from weighted sum of outputs in layer $n-1$:

$$x_j^{(n)} = f(\sum_i w_{i,j}^{(n)} x_i^{(n-1)} + w_{0,j}^{(n)})$$

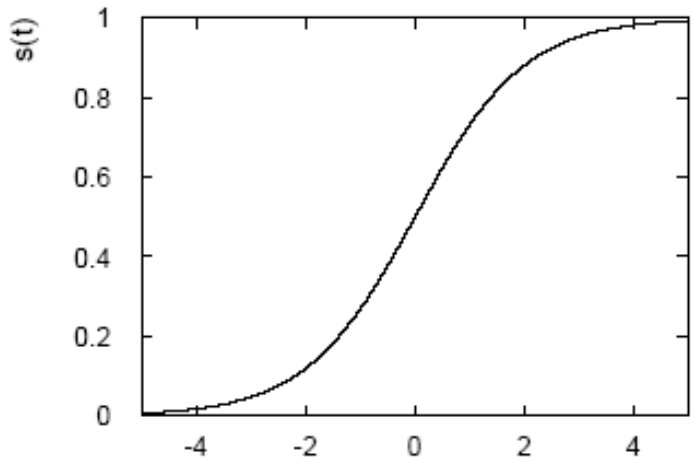
Each connection has associated a weight $w_{i,j}^{(n)}$, each node a bias $w_{0,j}^{(n)}$.

Neural network transfer functions

A non-linear monotonuous transfer function $f(x)$ is applied at the output of each node, e.g. the sigmoid function:

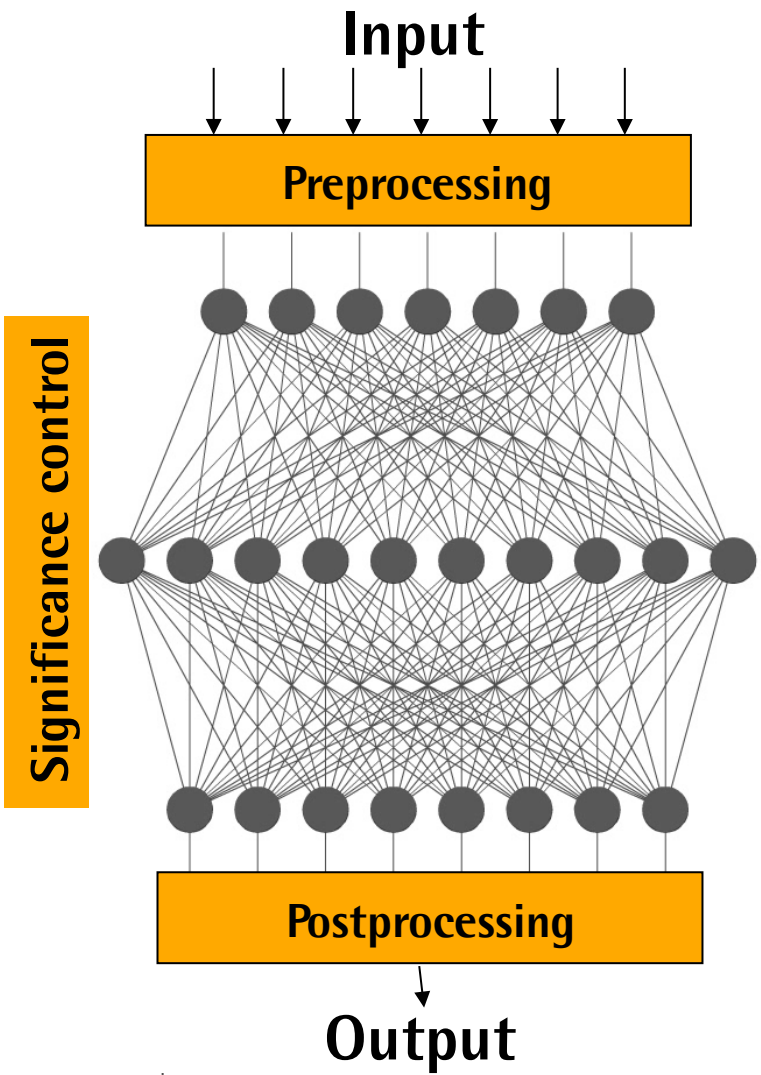
$$f(x) = \frac{1}{1 + exp(-x)}$$

It maps the intervall $(-\infty, \infty)$ to the compact $(0, 1)$.



NeuroBayes®

A very adavnced Neural Network – and much more



Invented in 2000 for reconstruction of b quark fragmentation in DELPHI experiment.

Further development in Phi-T,
later Blue Yonder.

Several hundred successful applications in DELPHI, CDF II, Belle, CMS, ATLAS, LHCb, H1, AMS experiments:
www.neurobayes.de

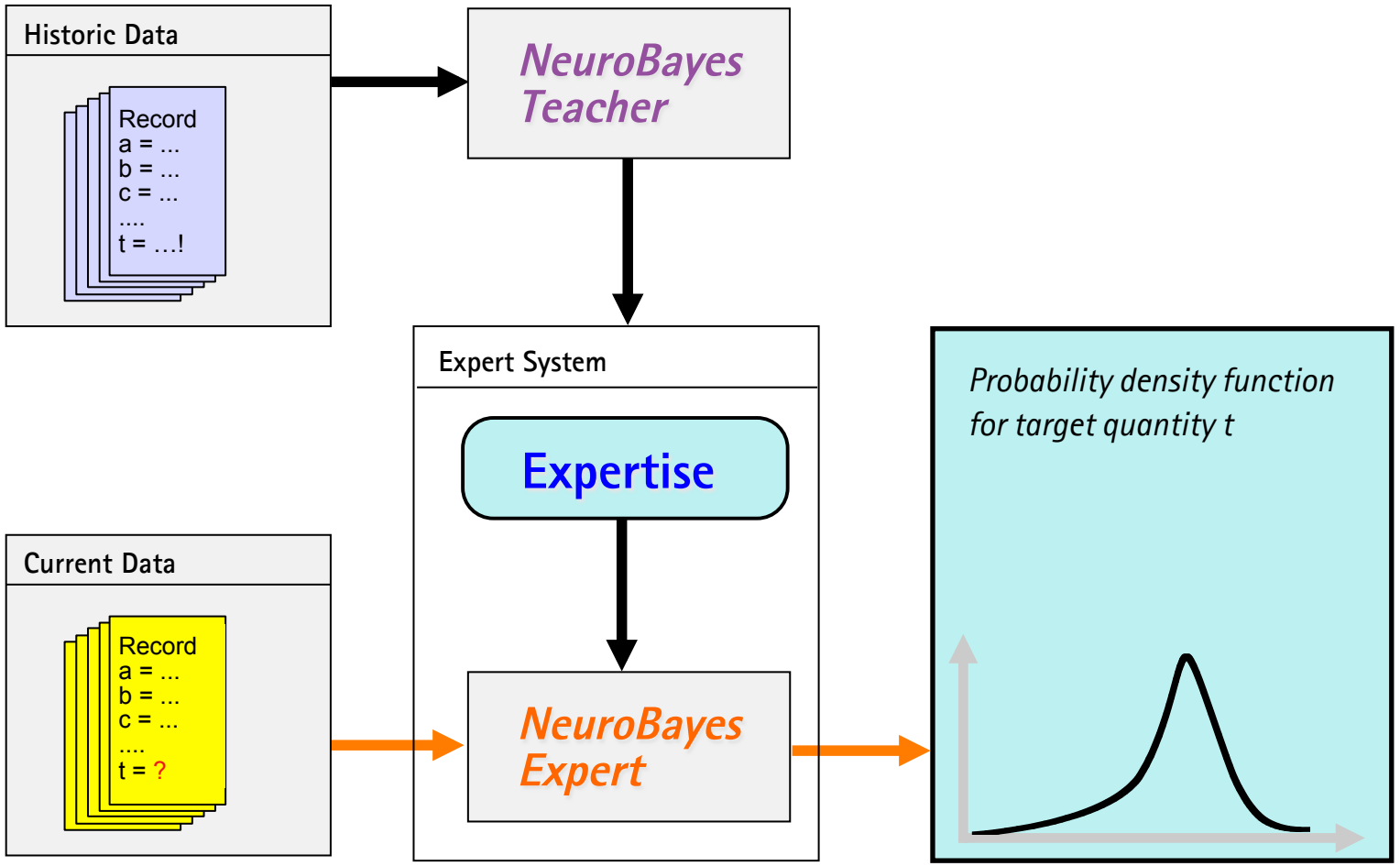
More than 400 men-years development.

Robust and fast algorithm for reconstruction (= prediction) of

- conditional probability densities
- classifications

with extreme generalization ability by means of Bayesian regularization.

NeuroBayes System Working principle



Neural network training

Training is the minimisation process of a loss function, during that the network weights are changed such that the deviation of the wanted output for a set of input vectors is minimisesd.

Possible loss functions:
Sum of quadratic deviations
or entropy (maximum likelihood)

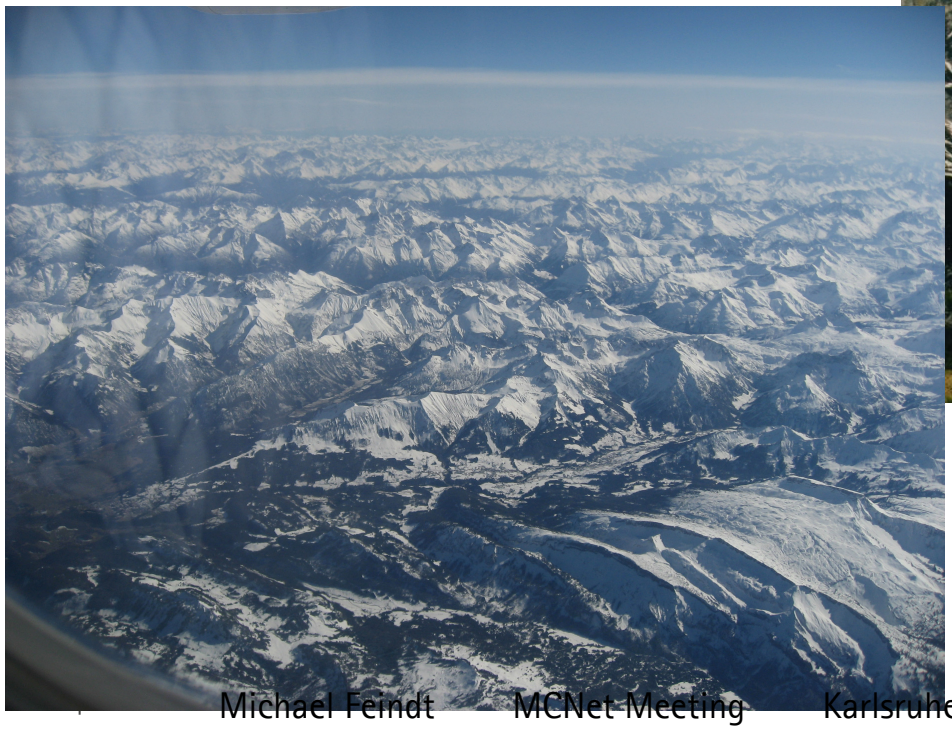
Backpropagation (Rumelhardt et al. 1986):
Calculate gradient backwards by applying chain rule
Optimise using gradient descent method. Step size??

Neural network training

Difficulty: find global minimum of highly non-linear function in high (~ >100) dimensional space.

Imagine task to find deepest valley in the Alps (just 2 dimensions)

Easy to find the next local minimum...



but globally...

...impossible!

➔ needs good preconditioning

NeuroBayes strengths:

NeuroBayes is a very powerful algorithm

- excellent generalisability (does not overtrain)
- robust – always finds good solution – even with erratic input data
- fast
- automatically selects significant variables
- output interpretable as Bayesian a posteriori probability
- can train with weights and background subtraction

NeuroBayes is easy to use

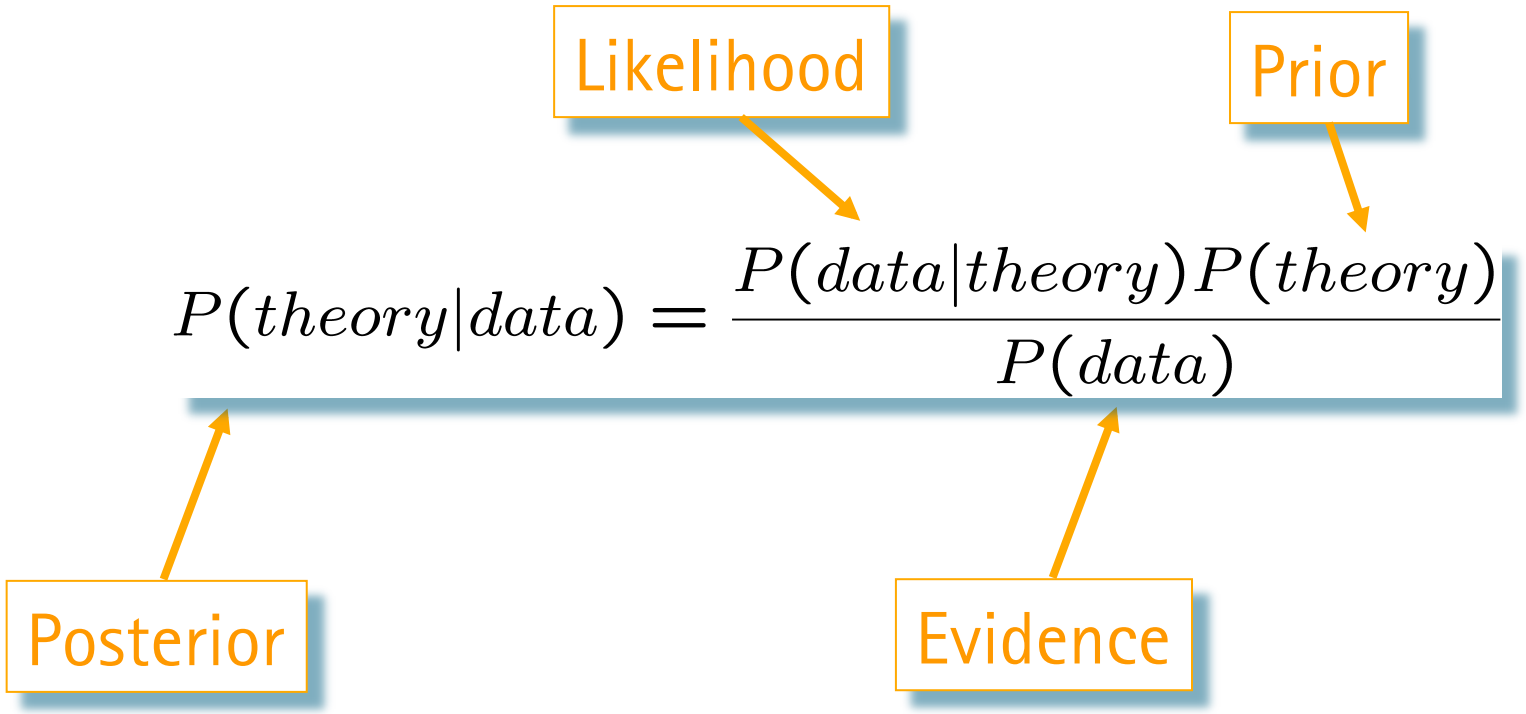
- Examples and documentation available
- Good default values for all options →fast start!
- Direct interface to root TMVA available

<phi-t> NeuroBayes®

> is based on 2nd generation neural network algorithms, Bayesian regularisation, optimised preprocessing with non-linear transformations and decorrelation of input variables and linear correlation to output.
> learns extremely fast due to 2nd order BFGS methods and even faster with 0-iteration mode.
> produces small expertise files.
> is extremely robust against outliers in input data.
> is immune against learning by heart statistical noise.
> tells you if there is nothing relevant to be learned.
> delivers sensible prognoses already with small statistics.
> can handle weighted events, even negative weights.
> has advanced boost and cross validation features.
> is steadily further developed professionally.

Bayes Theorem

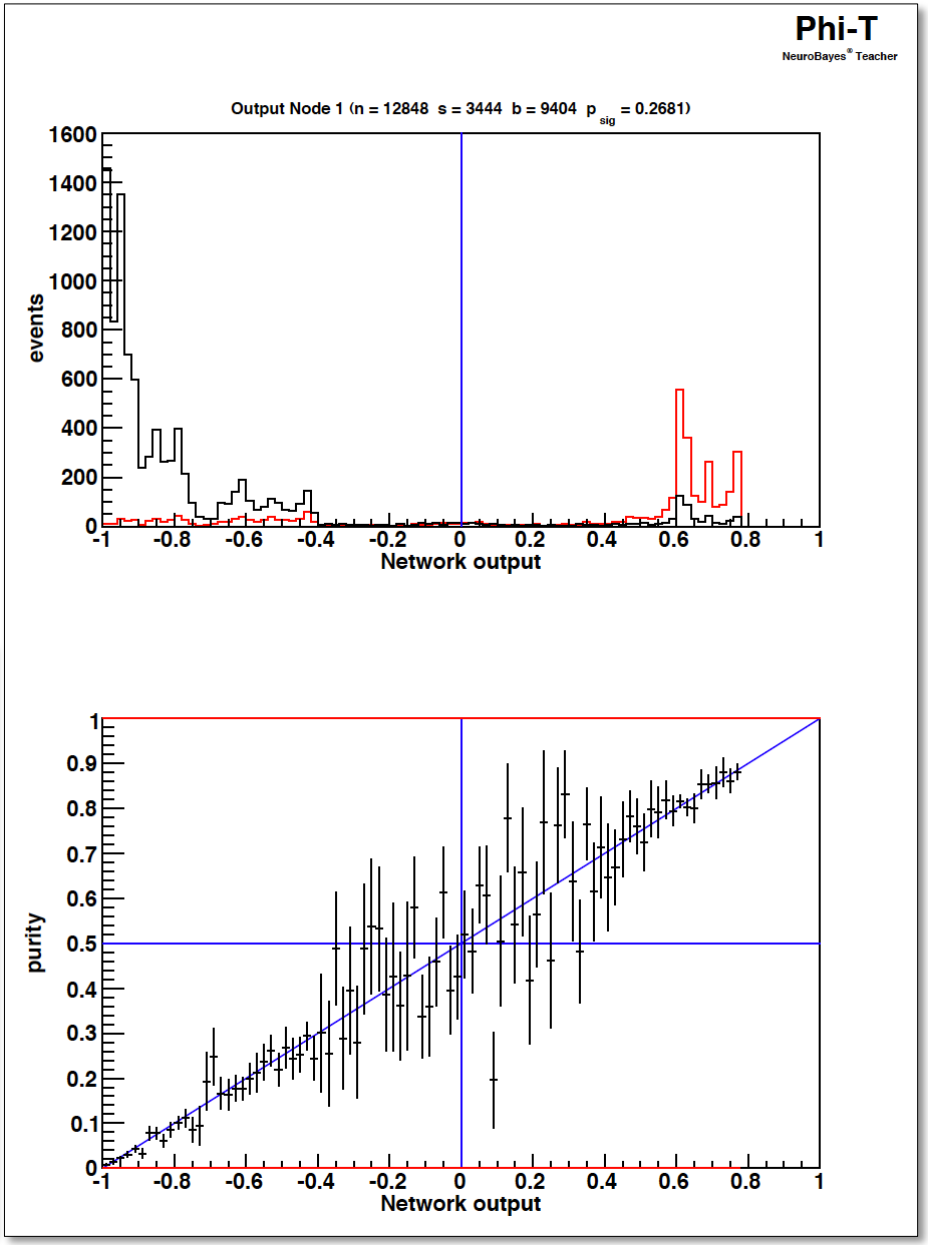
P(T | D) ≠ P(D | T) , but



NeuroBayes internally uses Bayesian arguments for regularisation
NeuroBayes automatically makes Bayesian posterior statements

NeuroBayes training output (analysis file)

NeuroBayes output distribution
red:signal black: background



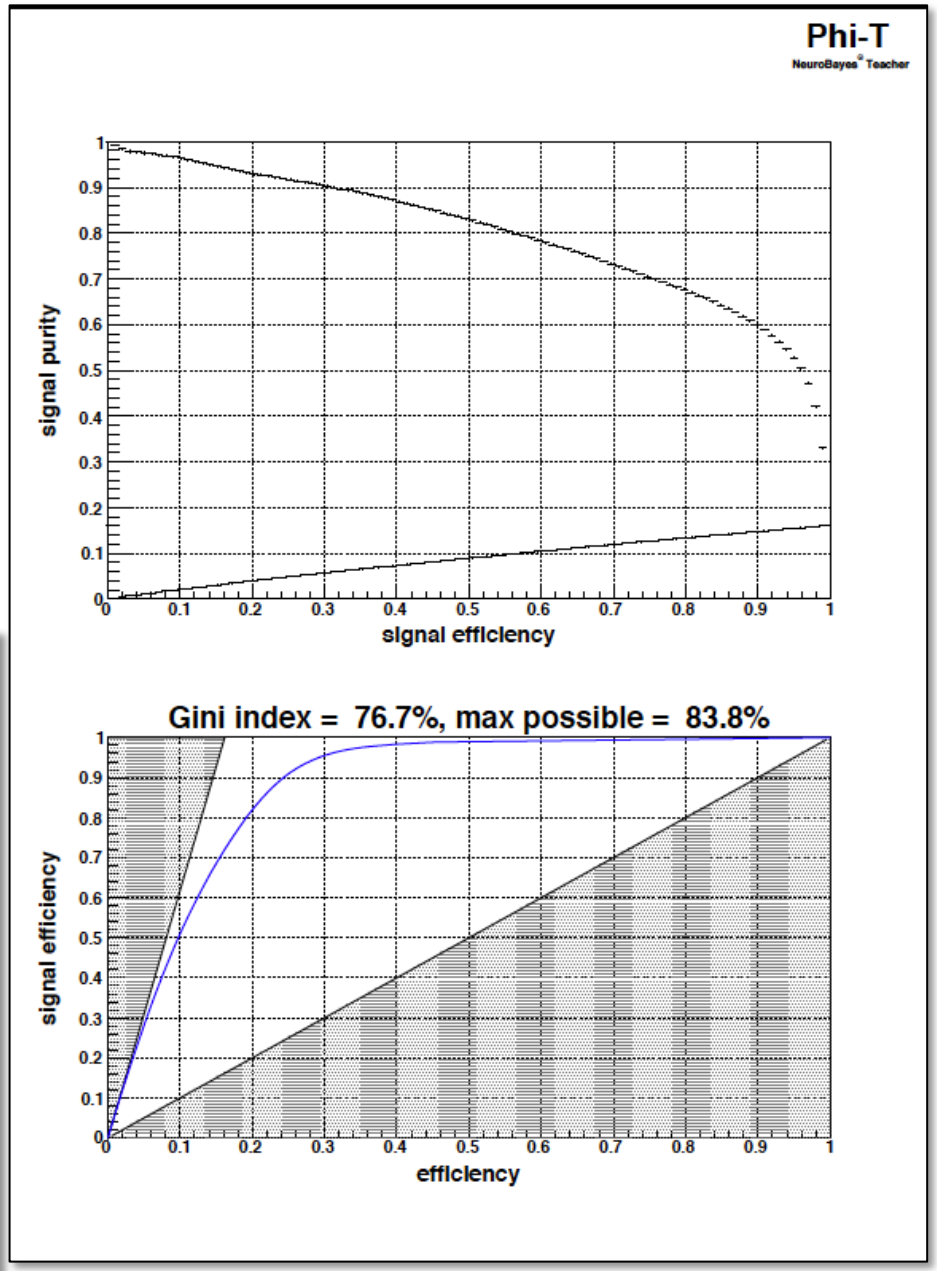
Signal purity S/(S+B) in bins of NeuroBayes output.

If on diagonal, then P=2*NBout+1 is the probability that the event actually is signal.
⇒This proves that NB always is well calibrated in the training.

NeuroBayes training output (analysis file)

Purity vs. signal efficiency plot for different NeuroBayes output cuts. Should be as much in upper right corner as possible.
The lower curve comes from cutting the wrong way round.

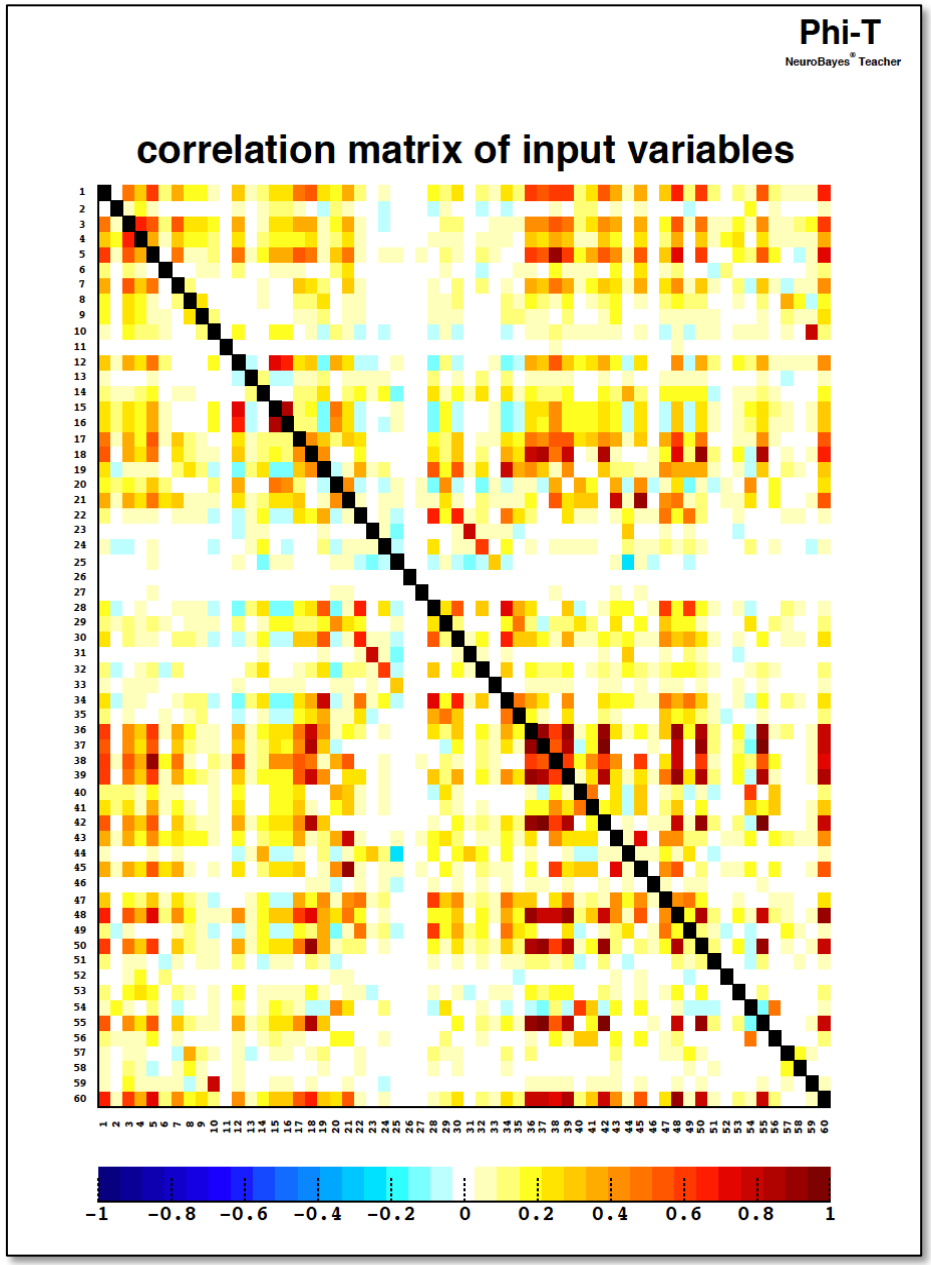
Signal efficiency vs. total efficiency when cutting at different NeuroBayes outputs (lift chart). The area between blue curve and diagonal should be large.
Physical region: white
Right diagonal: events randomly sorted, no individualisation.
Left diagonal border: completely correctly sorted, first all signal events, then all bg.
Gini index: classification quality measure, The larger, the better.



NeuroBayes training output (analysis file)

Correlation matrix of input variables.

1.row/column: training target



NeuroBayes training output (analysis file)

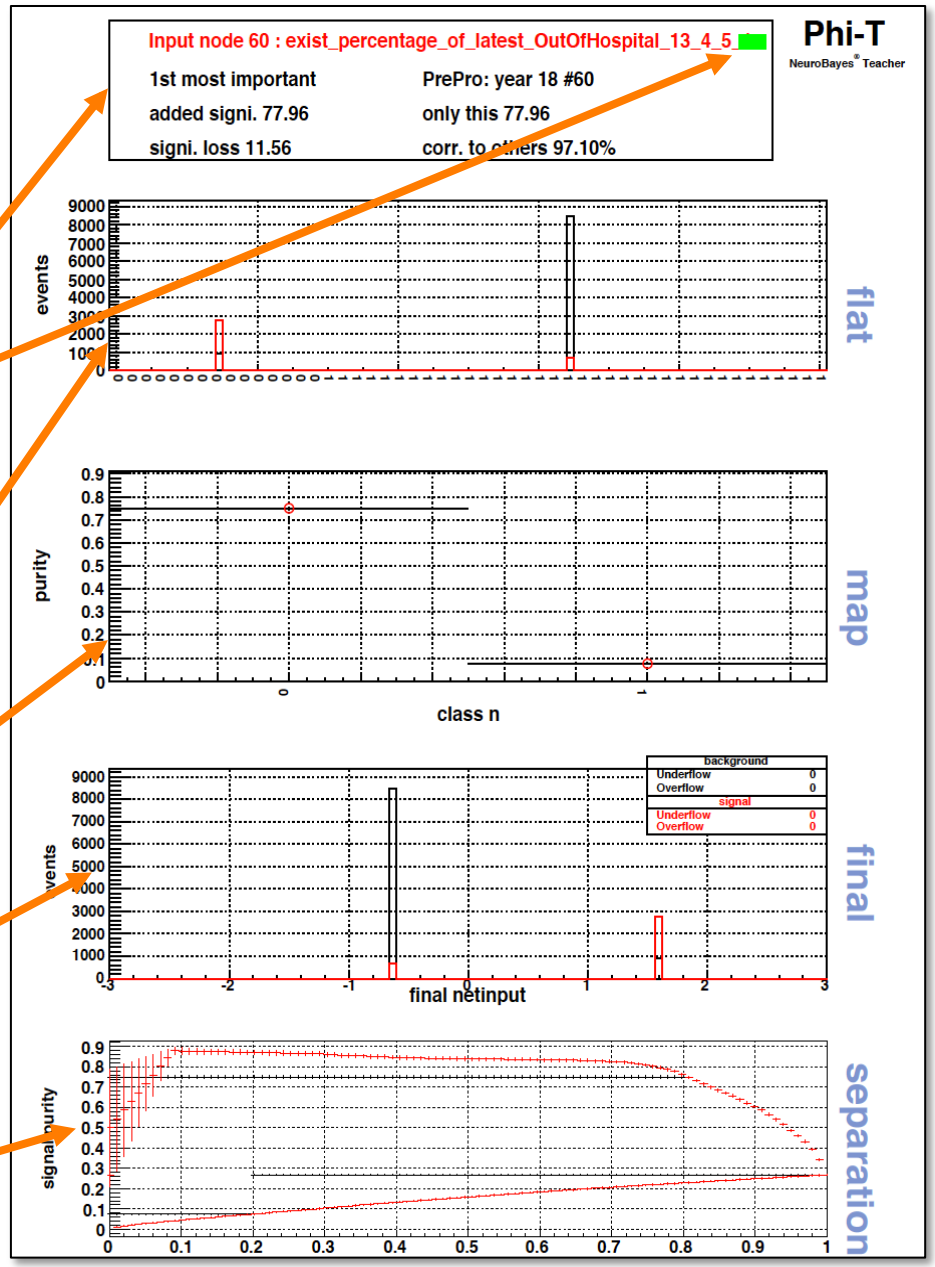
Most important input variable significance: 78 standard deviations
Accepted for the training

Probability integral transformed input variable distribution: signal, background (this is a binary variable!)

Signal purity as function of the input variable (this case: unordered classes)

Mean 0, width 1 transformation of signal purity of transformed input variable

Purity-efficiency plot of this variable compared to that of complete NeuroBayes



NeuroBayes training output (analysis file)

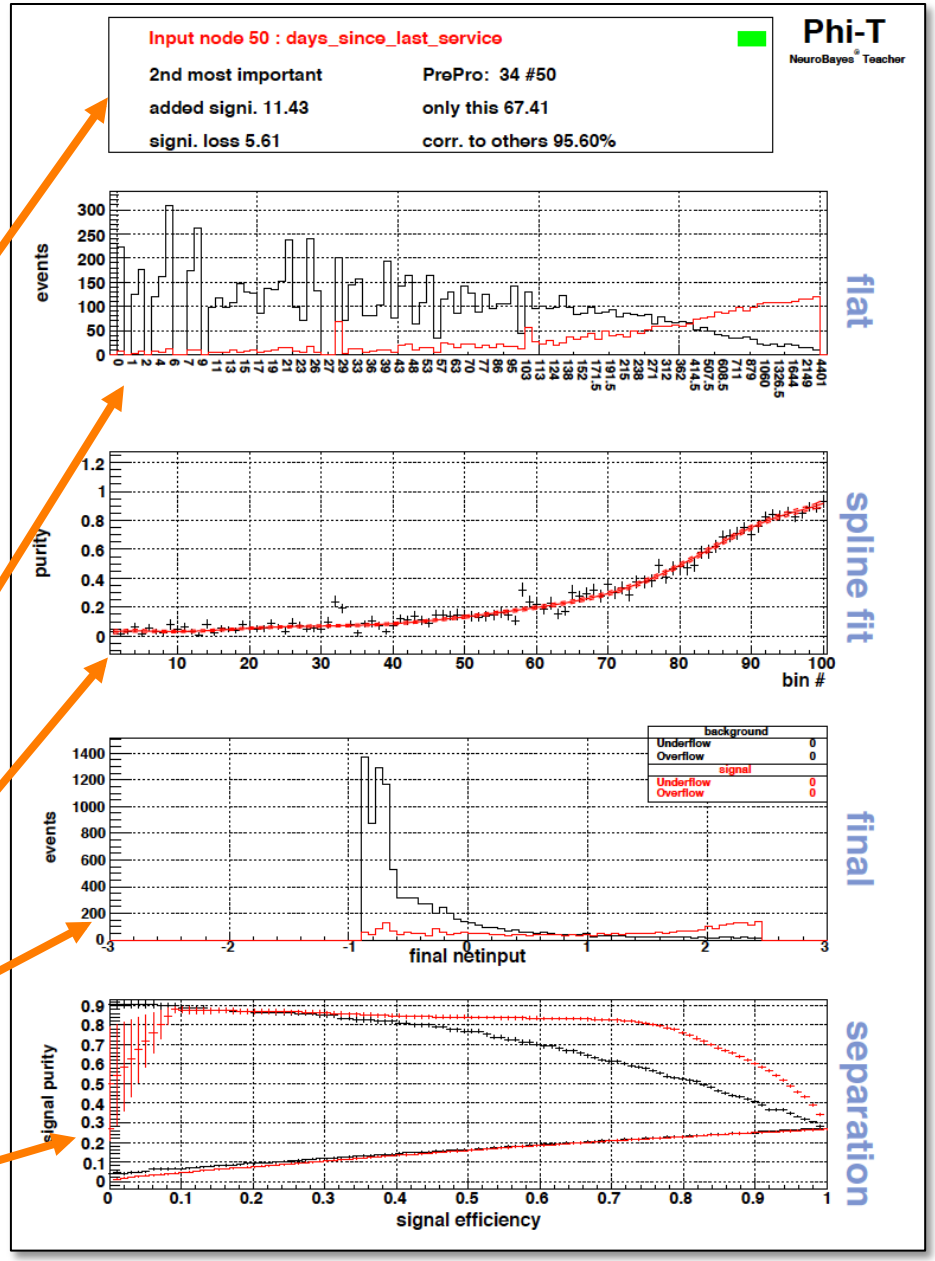
2.most important input variable, alone 67 standard deviations.
But added after most important var taken into account only 11 sigma.

Probability integral transformed input variable distribution: signal, background (this is a largely continuous variable!)

Signal purity as function of the input variable (this case: spline fit)

Mean 0, width 1 transformation of (fitted) signal purity of input variable

Purity-efficiency plot of this variable compared to that of complete NeuroBayes



NeuroBayes training output (analysis file)

39. most important input variable, alone 17 standard deviations, but only 0.6 sigma added after more significant variables.

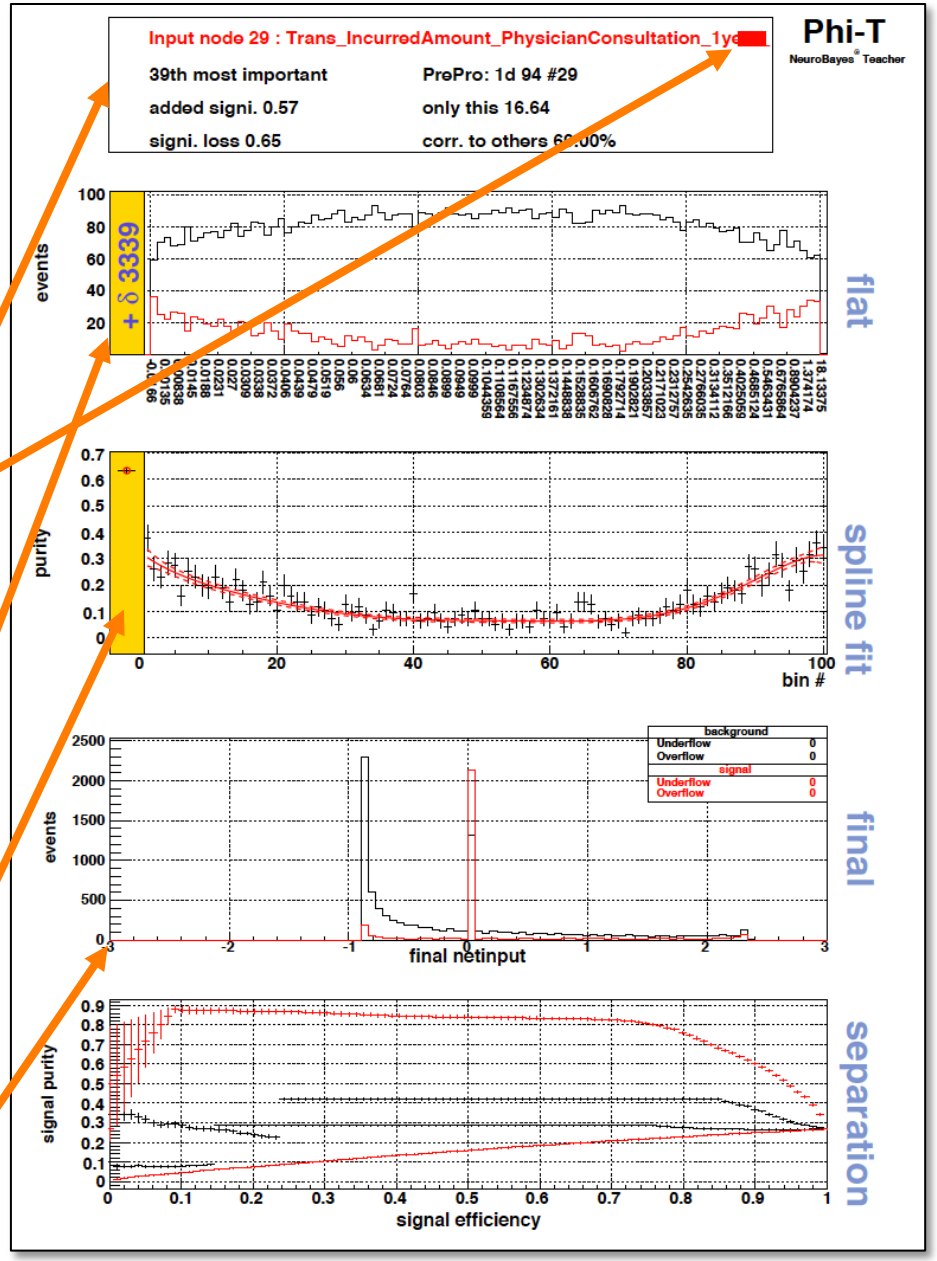
Ignored for the training

Probability integral transformed input variable distribution: signal, background

For 3339 events this input was not available (delta-function)

Signal purity as function of the input variable (this case: spline fit + delta)

Mean 0, width 1 transformation of (fitted) signal purity of input variable
Due to the preprocessing 94 the delta is mapped to 0, not to its purity.



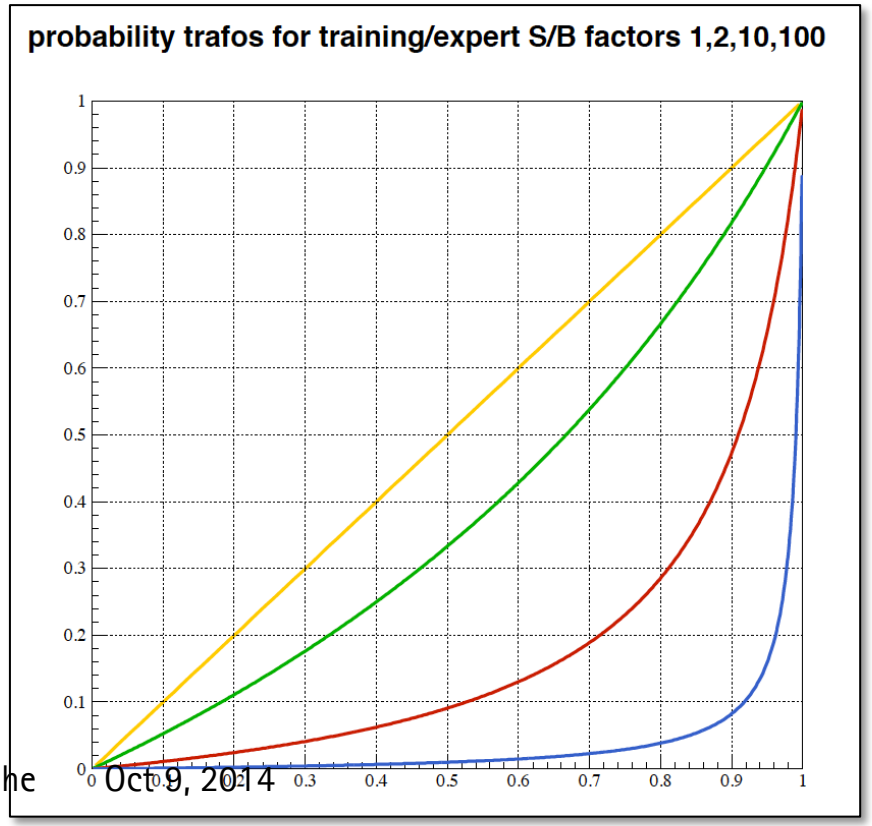
NeuroBayes output is a linear measure of the Bayesian posterior signal probability:

$$P_T(S) = (NB + 1)/2$$

Signal to background ratio in training set: $r_T = \frac{S_T}{B_T}$, in expert set: $r_E = \frac{S_E}{B_E}$

If the training was performed with different S/B than actually present in expert dataset, one can transform the signal probability:

$$P_E(S) = \frac{1}{1 + \left(\frac{1}{P_T(S)} - 1\right) \cdot \frac{r_T}{r_E}}$$



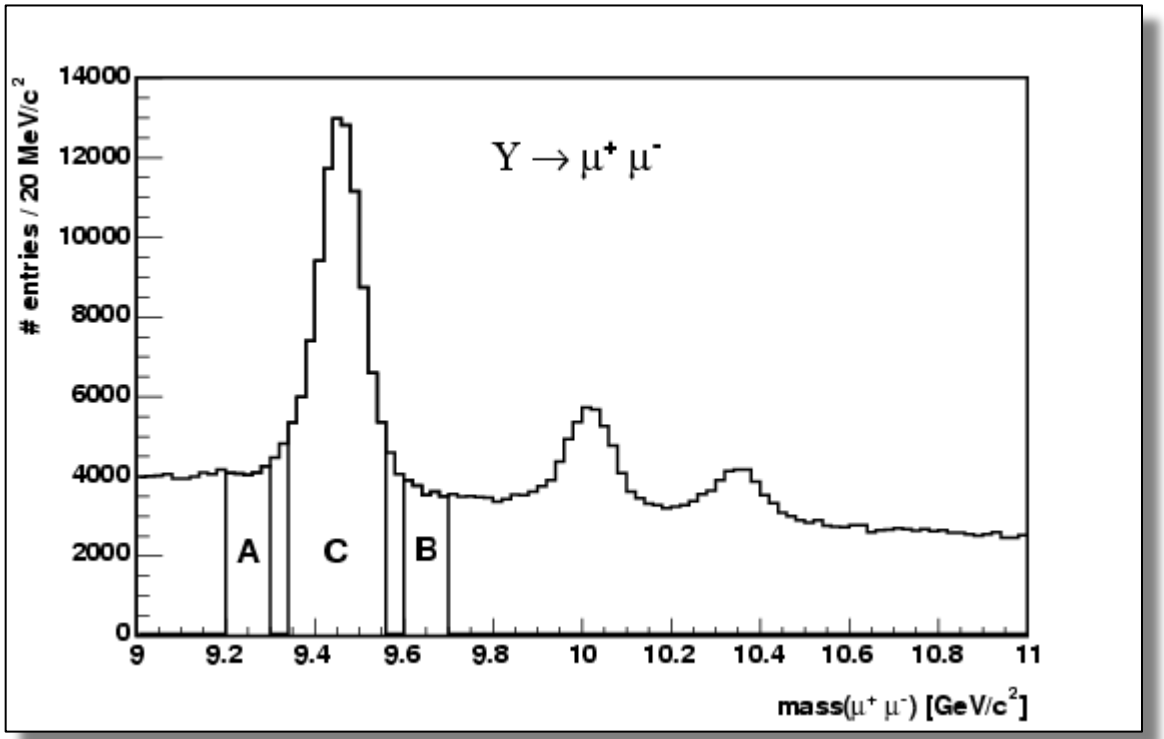
Scenario: Neither reliable signal nor background Monte Carlo available

Idea: Training with background subtraction

Signal: Peak region weight 1

Sideband region with weight -1 (statistical subtraction)

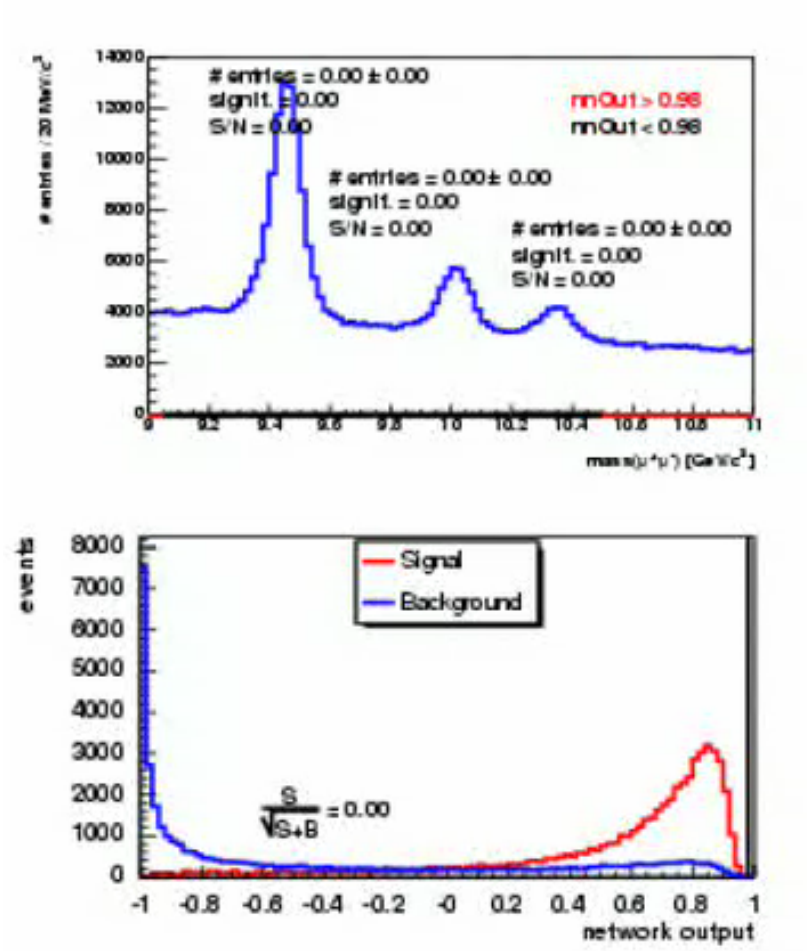
Background: Sideband region with weight 1

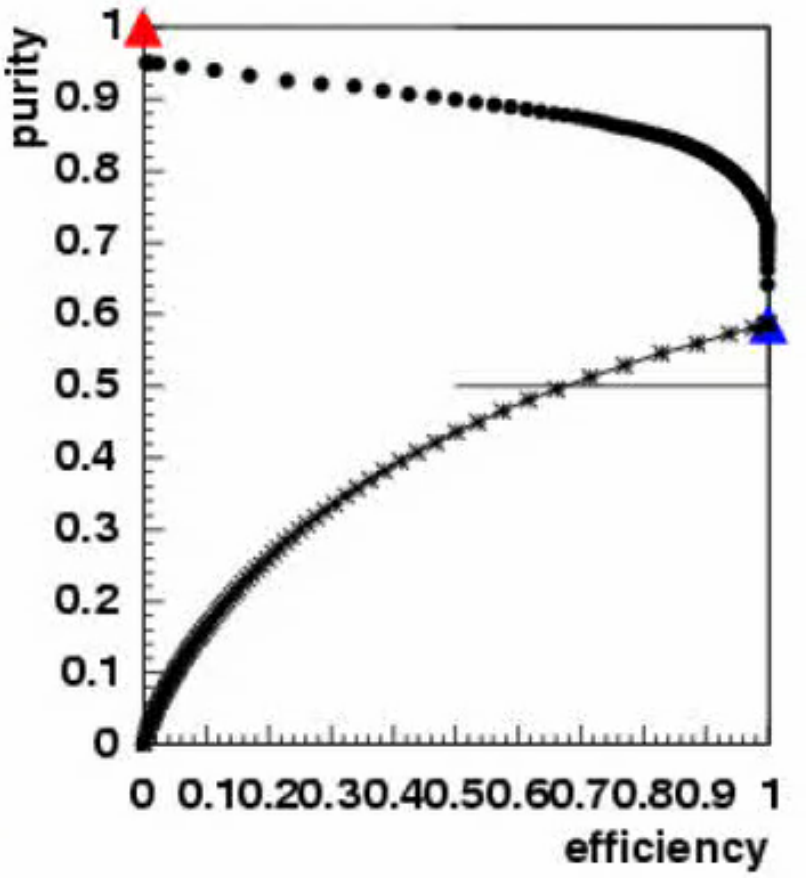


works very well!

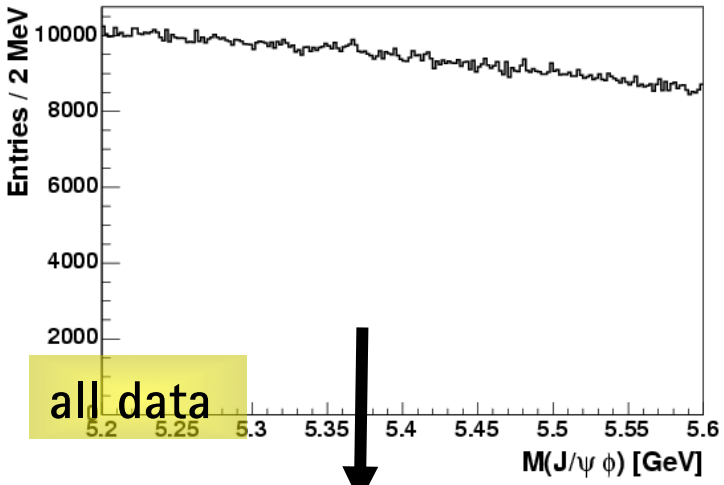
also for Y(2S) and Y(3S) ! Although just trained on Y(1S)

Example for data-only training (on 1. resonance)

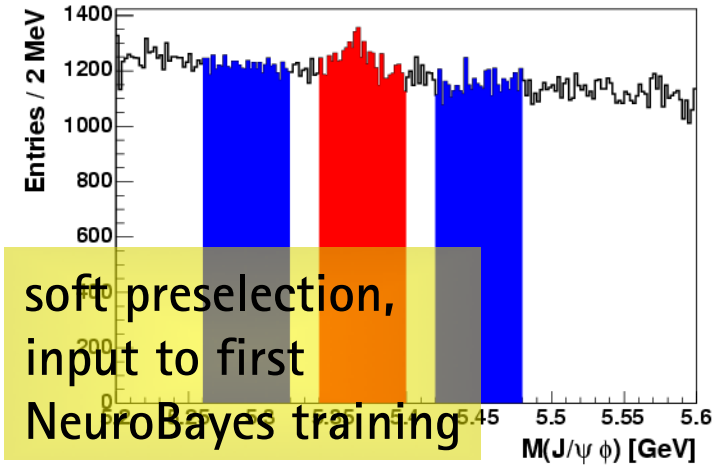




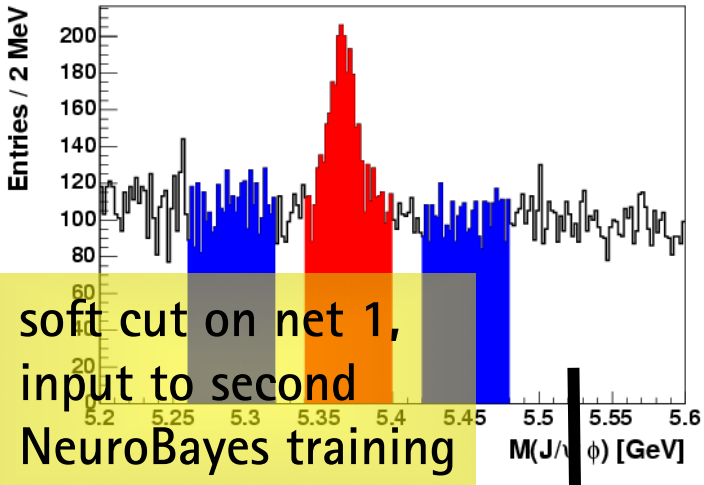
NeuroBayes B_s to $J/\psi\ \Phi$ selection without MC (2 stage background subtraction training process)



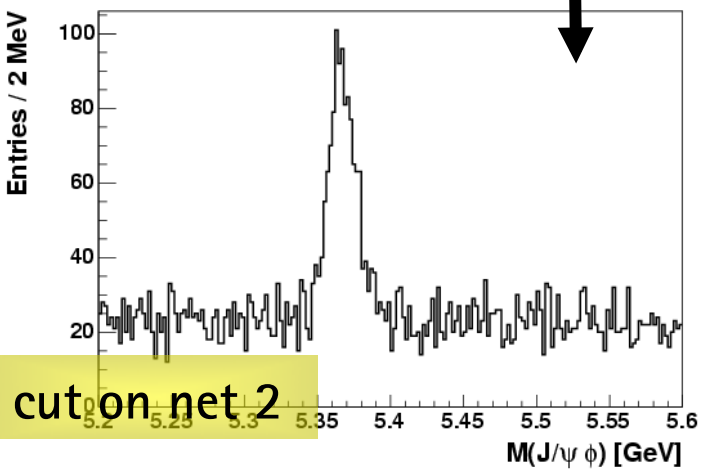
all data



soft preselection, input to first NeuroBayes training



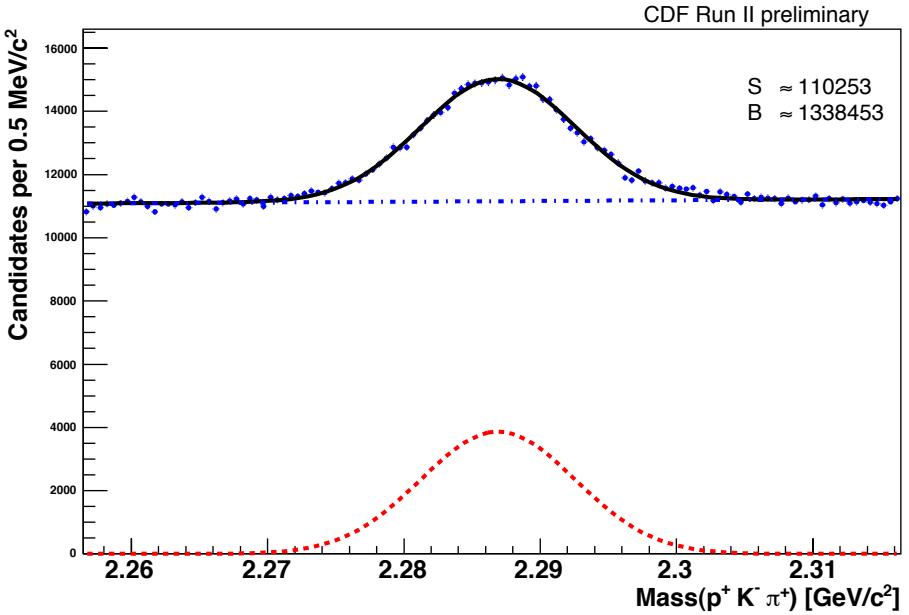
soft cut on net 1, input to second NeuroBayes training



cut on net 2

Exploiting S/B information more efficiently : The sPlot-method

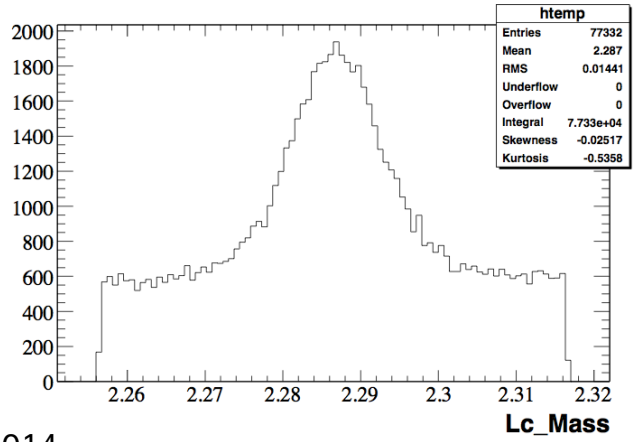
Fit data signal and background in one distribution (e.g. mass). Compute sPlot weights w_s for signal (may be <0 or >1) as function of mass from fit.



Train NeuroBayes network with each event treated both as signal with signal weight w_S and as background with weight $1-w_S$.

Soft cut on output enriches S/B considerably:

Make sure network cannot learn mass!



More than 120 Ph.D. theses and many publications ...

from experiments DELPHI, CDF II, AMS, CMS ATLAS, LHCb and Belle used NeuroBayes® or predecessors very successfully.

Many of these can be found at
www.neurobayes.de

Talks about NeuroBayes® and applications:
www-ekp.physik.uni-karlsruhe.de/~feindt → Forschung

Some NeuroBayes highlights: Bs oscillations

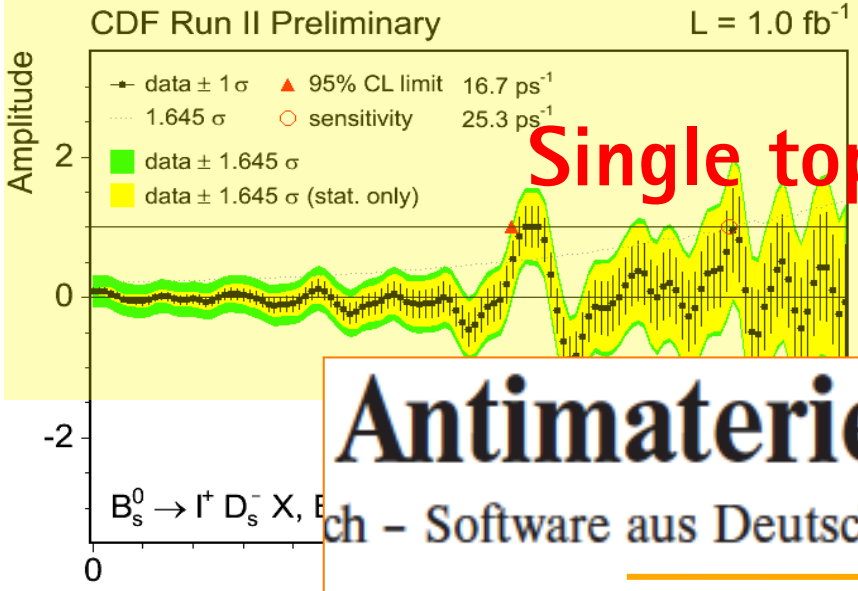
Discovery of excited Bs states

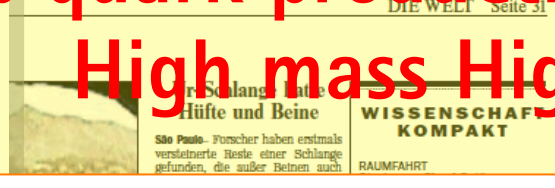
X(3872) properties

Single top quark production discovery

High mass Higgs exclusion

...





DIE WELT Seite 31

...r Schlange fand Hüfte und Beine

São Paulo – Forscher haben erstmals versteinerte Reste einer Schlange gefunden, die außer Beinen auch

WISSENSCHAFT KOMPAKT

RAUMFAHRT

Antimaterie wandeln

...ch – Software aus Deutschland ermöglichte die Messungen

...Sekunde 2,8 Bil-
...-B-Mesonen um
...ede Sekunde et-

Tempo dieses Tanzes gemessen", sagt Jacobo Konigsberg, Sprecher der CDF-Kollaboration.

kenntnisse über die Eigenschaften der Elementarteilchen, sondern auch über die Entwicklung des

von Medikamenten zu befassen. Der Arzt schrieb auf, was helfen sollte, und das besorgte man sich im „krüdhus", für das sich schon im 13. Jahrhundert die Bezeichnung Apotheke einbürgerte. Das aus dem Griechischen entlehnte Wort bedeutet „Magazin", ein „Ort, wo man etwas aufbewahrt".

In der Überlieferung der Familiennamen geschieht aber Seltsames. Während in den alten Urkunden die Berufsbezeichnung Apothekarius, eingedeutscht zu Appentegger und Apteker, auch als Familienname oft vorkommt, gibt es ihn heute praktisch nicht mehr. Dafür ist an deutschen Apothekernamen kein Mangel: Apotheker waren die Vorfahren des Komponisten Peter Kreuder. Dazu gehört eine riesige Namenfamilie: Kräuter, Kreuter, Kreutler, Kräutler, Kreußler, Krautner, Kräutner, Krude, Krüder und Krüdener. Nicht zu vergessen die Wurzler, Würzler, Würzner, Wurzer, Wurz und Wurtz. Das bekannteste heimische heilkräftige Kraut, das Origanum vulgare, auf deutsch: Dost, schenkte die Namen Doster, Dostler und Dostmann. Der Salbenhersteller darf nicht fehlen: Schmerschneider. *Hans Markus Thomsen*

Das unvollständige Skelett der Ur-Schlange Najash rionegrina wurde in Argentinien entdeckt. Es bedeutet für Paläobiologen eine Sensation. Die Wissenschaftler fanden bei dem sehr ursprünglichen Tier Knochen von kräftigen Beinen, ein Kreuzbein und einen Beckengürtel FOTO: AP

Riesenschlangen gibt es Reste von Becken und Oberschenkel. *DW*

seit vielen Jahren bekannt, doch erst jetzt entdeckten die Forscher, daß es bei der Aktivierung von genetischen Programmen im Zellkern eine zentrale Rolle spielt und folglich auch an der Entstehung von Fehlfunktionen ...rt beteiligt sein kann. *AP*

Glückshormon gut für die Leber
Das Glückshormon Serotonin fördert die Regeneration von verletztem Lebergewebe. Dies beobachteten Forscher des Max-Planck-Instituts für molekulare Genetik und des Max-Delbrück-Centrums in Berlin. *N.L.*

Musik lindert Schmerzen
Patienten, die während einer Operation in örtlicher Betäubung Musik hören, benötigen weniger Schmerzmittel, berichtet die Ärztezeitschrift „Praxis-Depesche". *ls.*

Materie kann sich in Antimaterie wandeln

Amerikanische Elementarteilch...nphysiker melden Durchbruch – Software aus Deutschland ermöglichte die Messungen

Von Christian Meier

Chicago – Elementarteilchenphysiker melden eine Sensation: Sie haben erstmals die Umwandlungen zwischen Materie und Antimaterie direkt beobachtet. Das seit vielen Jahren existierende Standardmodell der Teilchenphysik – also das vorherrschende Modell für die kleinsten Teilchen und der Kräfte zwischen ihnen – sagt voraus, daß so genannten B-Mesonen die einzigartige Fähigkeit besitzen, sich spontan in ihr Antiteilchen umwandeln zu können – und umgekehrt. Jetzt ist es US-Physikern am Fermilab bei Chicago gelungen, die extrem schnelle Umwandlung zeitlich aufgelöst zu beobachten und damit die theoretische Vorhersage experimentell zu bestätigen.

Als einzige deutsche ...titution war die Universität Karlsruhe maßgeblich an dem Experiment beteiligt. Zwanzig Physiker um Thomas Müller und Michael Feindt haben die komplexe Software für eine gezielte Auswertung der Rohdaten geliefert. Das Team gehört zu der Kollaboration „Collider Detector at Fermilab" (CDF), an der etwa 700 Physiker von 60 Institutionen beteiligt sind.

Im Fermilab, dem leistungsfähigsten Teilchenbeschleuniger der Welt, werden Protonen und Antiprotonen auf nahezu Lichtgeschwindigkeit beschleunigt und dann aufeinander geschossen. Die dabei neben vielen anderen Teilchen entstehenden B-Mesonen wandeln sich pro Sekunde 2,8 Billionen Mal in Anti-B-Mesonen um ...nd zurück, also jede Sekunde etwa 500-mal...e Menschen auf der Erde leben. „Dieser Wert liegt im Bereich, den das Standardmodell vorhersagt", erläutert Müller.

„Wenn man sich Materie vorstellt, die mit Antimaterie tanzt, dann haben wir das unglaubliche Tempo dieses Tanzes gemessen"

Jacobo Konigsberg, Sprecher der Teilchenphysiker-Gruppe

„Wenn man sich Materie vorstellt, die mit Antimaterie tanzt, dann haben wir das unglaubliche Tempo dieses Tanzes gemessen", sagt Jacobo Konigsberg, Sprecher der CDF-Kollaboration.

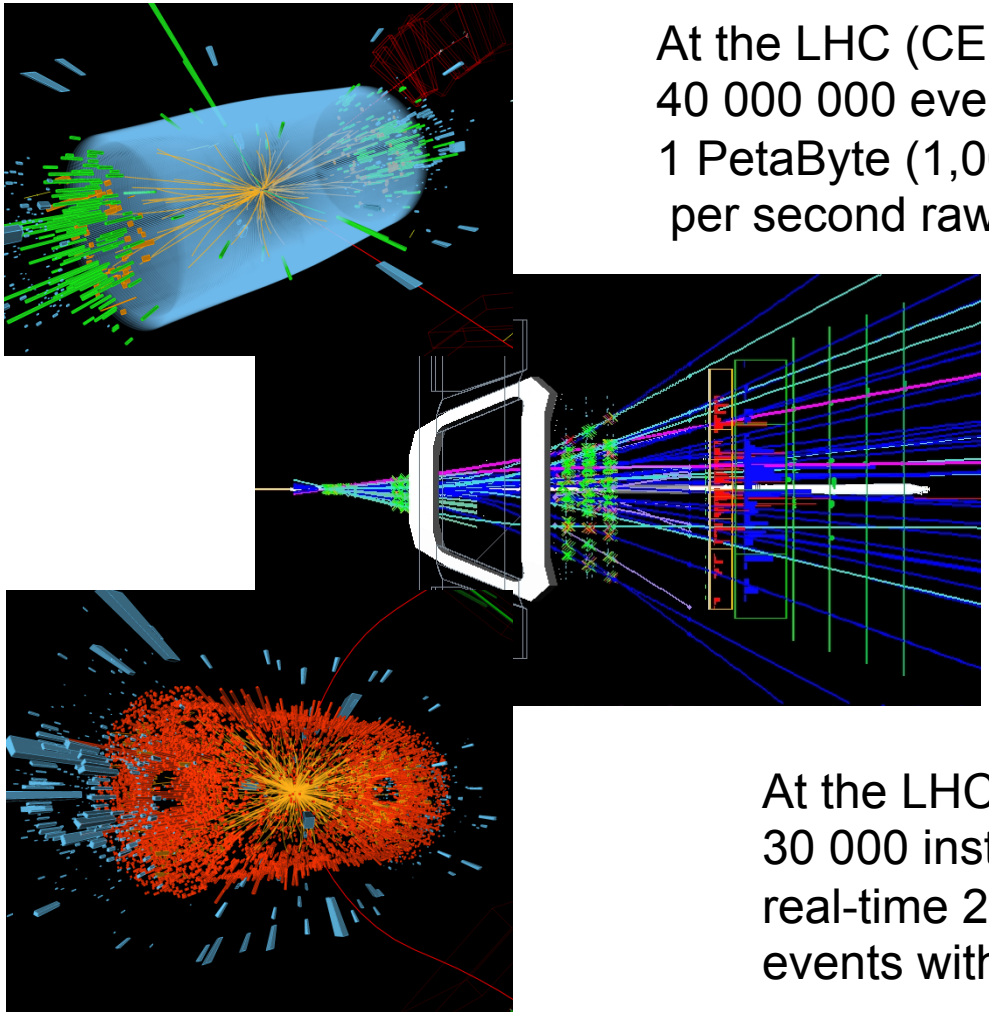
B-Mesonen existieren im heutigen Kosmos nicht mehr, waren aber im jungen Universum kurz nach dem Urknall vorhanden. Physiker können sie nur in großen Teilchenbeschleunigern untersuchen. Sie wollen so nicht nur Erkenntnisse über die Eigenschaften der Elementarteilchen, sondern auch über die Entwicklung des ...rsums gewinnen.

Seit 1995 arbeiteten die Karlsruher an Software, die aus dem Gewirr elektronischer Teilchenspuren im CDF-Detektor rekonstruieren kann, ob ein B-Meson bei seiner Entstehung Teilchen oder Antiteilchen war. Dies gelang mit komplexen statistischen Verfahren. Zusammen mit der Messung der Lebensdauer des B-Mesons (rund eine Millionstel Millionstel Sekunde) und der relativ einfach zu gewinnenden Information, ob es beim seinem Zerfall Teilchen oder Antiteilchen war, kann auf die Anzahl der Umwandlungen pro Sekunde geschlossen werden.

Das Ressort Wissenschaft erreichen Sie unter:
Telefon: 030 25 91 - 7 19 68
Fax: 030 25 91 - 7 19 67
E-Mail: wissenschaft@welt.de
Internet: www.welt.de/wissenschaft

NeuroBayes example: The LHCb trigger

very fast intelligent decisions with NeuroBayes



At the LHC (CERN) – per experiment:
40 000 000 events per second, which translates into
1 PetaByte (1,000,000,000,000,000 Byte)
per second raw data

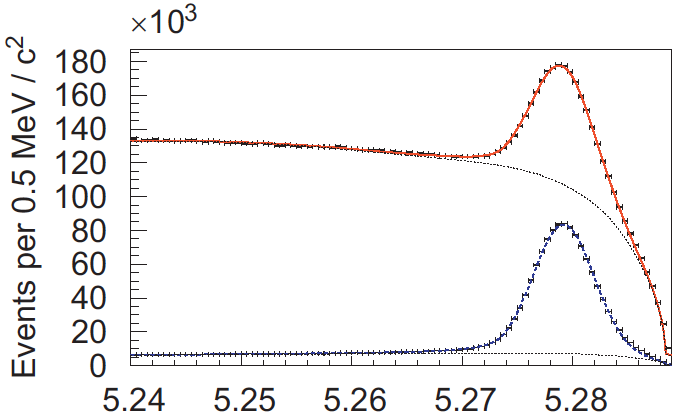
But only 1 PB of interesting data per year can be stored.

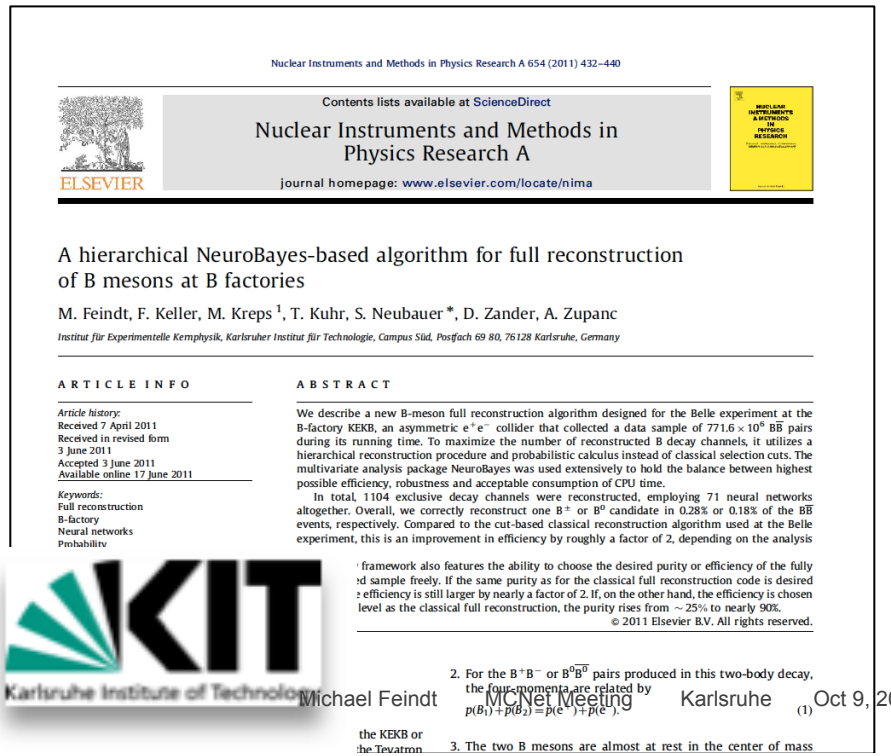
Need online reduction by
1 : 10,000,000

At the LHCb experiment
30 000 instances of NeuroBayes running
real-time 24/7 filter out the „interesting“
events without introducing lifetime bias

Photo: CERN

NeuroBayes example: Full reconstruction of B mesons at the Japanese B factory experiment Belle





Nuclear Instruments and Methods in Physics Research A 654 (2011) 432–440

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

A hierarchical NeuroBayes-based algorithm for full reconstruction of B mesons at B factories

M. Feindt, F. Keller, M. Kreps [1], T. Kuhr, S. Neubauer *, D. Zander, A. Zupanc

Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, Campus Süd, Postfach 69 80, 76128 Karlsruhe, Germany

ARTICLE INFO

Article history:
Received 7 April 2011
Received in revised form
3 June 2011
Accepted 3 June 2011
Available online 17 June 2011

Keywords:
Full reconstruction
B-factory
Neural networks
Probability

ABSTRACT

We describe a new B-meson full reconstruction algorithm designed for the Belle experiment at the B-factory KEKB, an asymmetric e^+e^- collider that collected a data sample of 771.6×10^6 $B\overline{B}$ pairs during its running time. To maximize the number of reconstructed B decay channels, it utilizes a hierarchical reconstruction procedure and probabilistic calculus instead of classical selection cuts. The multivariate analysis package NeuroBayes was used extensively to hold the balance between highest possible efficiency, robustness and acceptable consumption of CPU time.

In total, 1104 exclusive decay channels were reconstructed, employing 71 neural networks altogether. Overall, we correctly reconstruct one B^+ or B^0 candidate in 0.28% or 0.18% of the $B\overline{B}$ events, respectively. Compared to the cut-based classical reconstruction algorithm used at the Belle experiment, this is an improvement in efficiency by roughly a factor of 2, depending on the analysis

framework also features the ability to choose the desired purity or efficiency of the fully
d sample freely. If the same purity as for the classical full reconstruction code is desired
e efficiency is still larger by nearly a factor of 2. If, on the other hand, the efficiency is chosen
level as the classical full reconstruction, the purity rises from ~25% to nearly 90%.

2. For the B^+B^- or $B^0\overline{B^0}$ pairs produced in this two-body decay, the four-momenta are related by $p(B_1)+p(B_2)=p(e^-)+p(e^+)$. (1)

the KEKB or
the Tevatron

3. The two B mesons are almost at rest in the center of mass

- Belle experiment at KEK/Japan
- 400 physicists from whole world
- 10 years of data taking and analysis
- World record luminosity
- > 400 publications

- Automatic hierarchical reconstruction system built from 72 NeuroBayes networks reconstructed about 1100 different reactions with a factor 2 larger efficiency than all analyses before

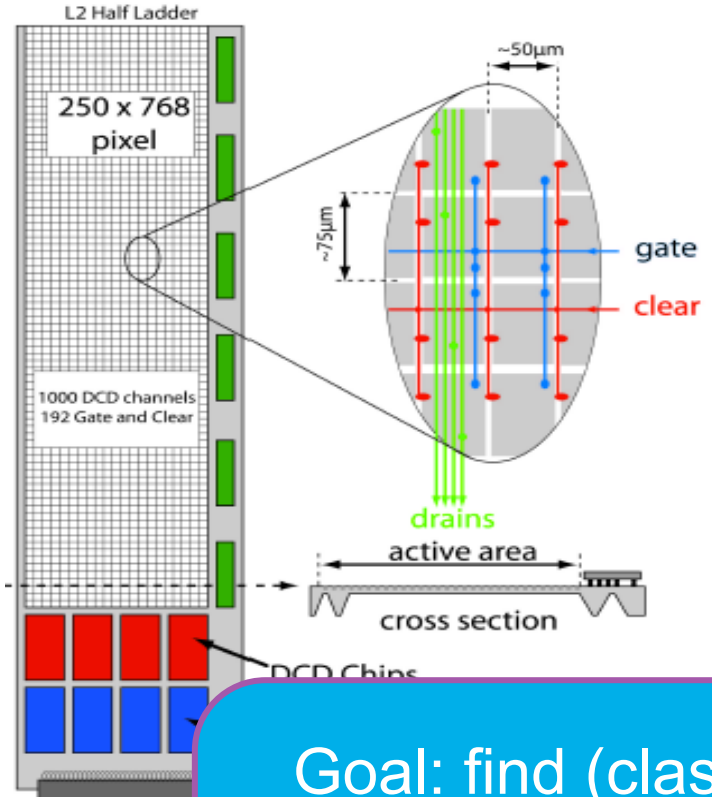
- Much cleaner signal

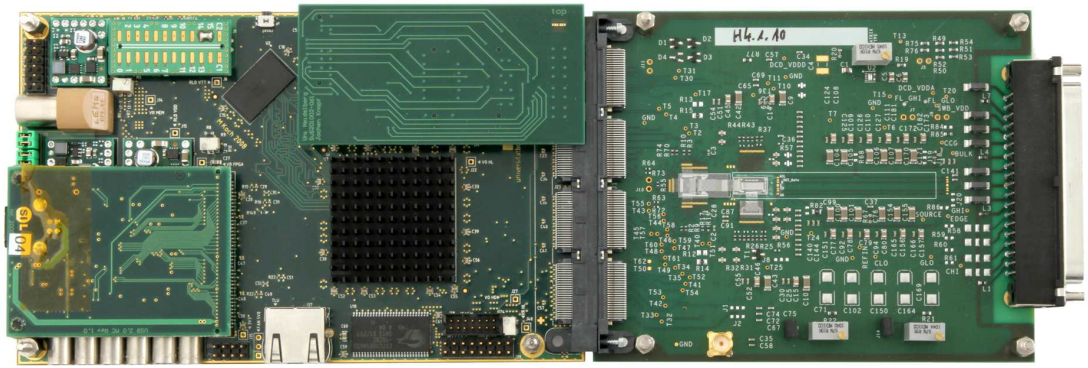
- Work performed by 2 PhD and 1 master student
- Corresponds to 500 “normal” PhD theses
- Corresponds to another 10 years of data taking

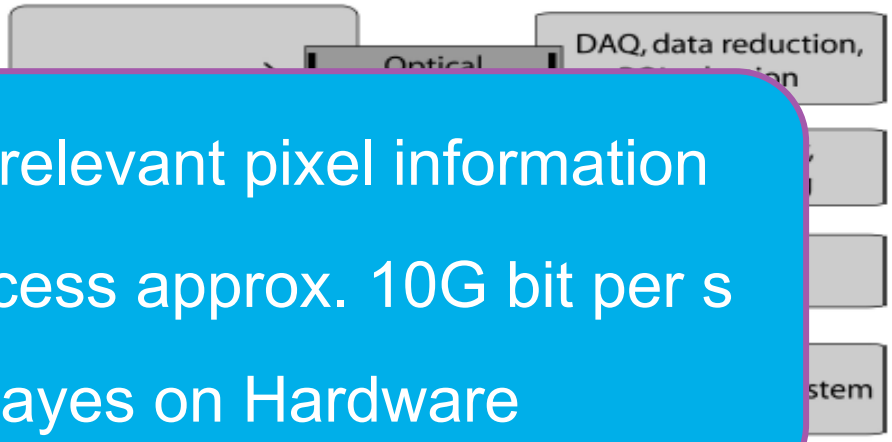
blue yonder

Forward looking. Forward thinking.

Future (2015): intelligent decisions directly on sensor (Belle II pixel detector), before big data reaches any computer







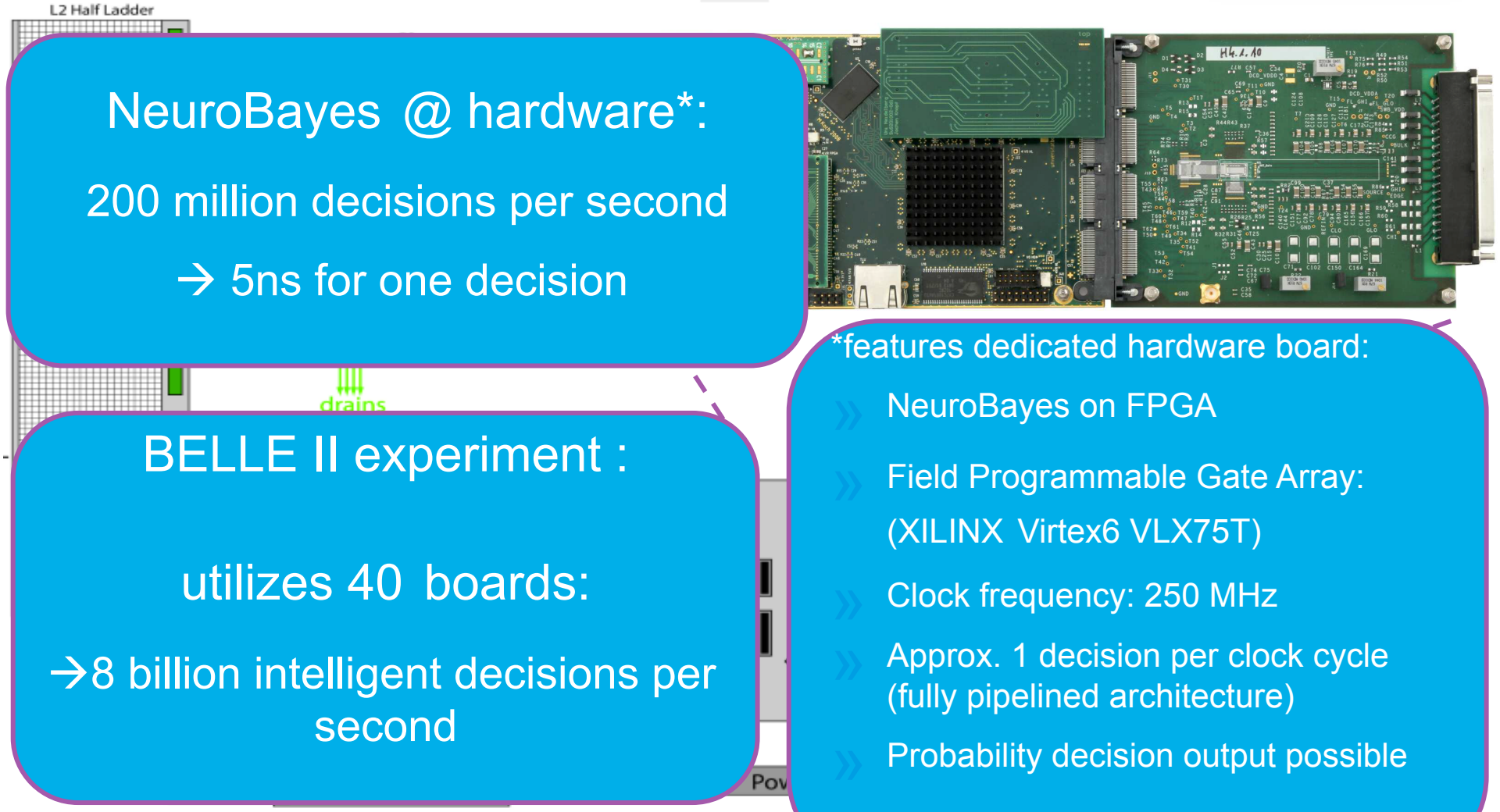
Goal: find (classify) all relevant pixel information

Big Data challenge: process approx. 10G bit per s

Solution: NeuroBayes on Hardware

blue yonder
Forward looking. Forward thinking.

KIT
Karlsruhe Institute of Technology



NeuroBayes @ hardware*:

200 million decisions per second

→ 5ns for one decision

BELLE II experiment :

utilizes 40 boards:

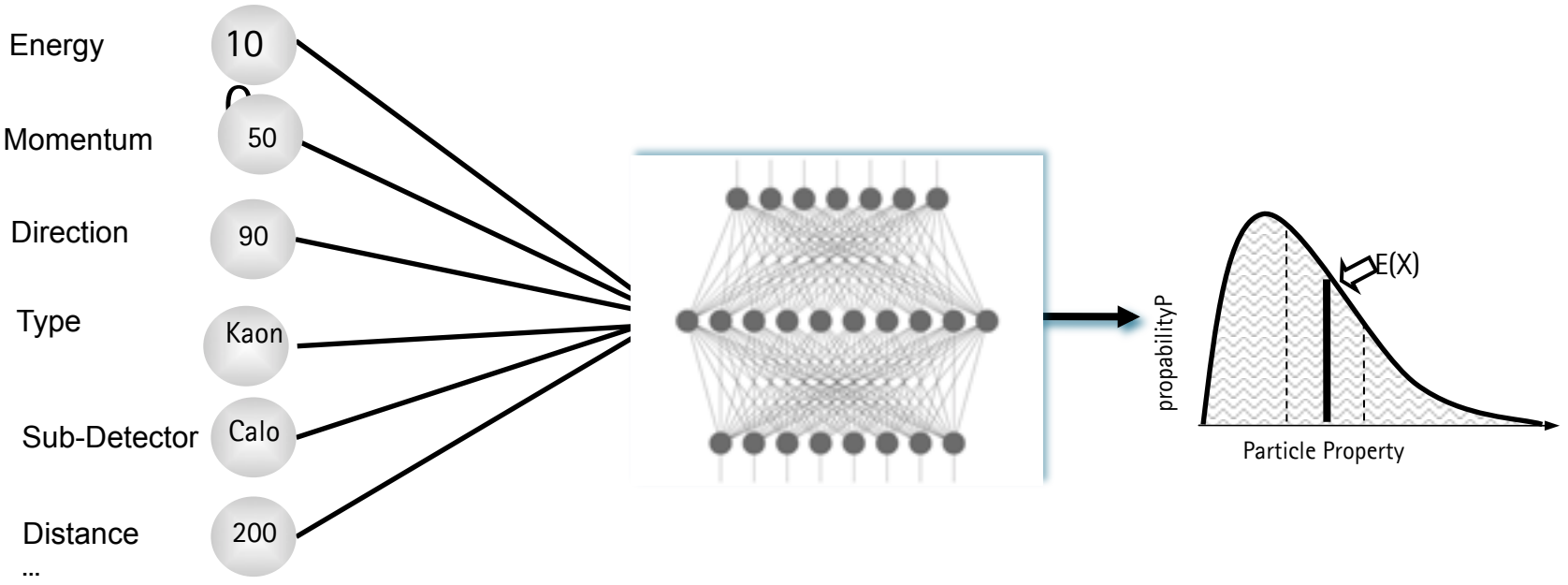
→8 billion intelligent decisions per second

*features dedicated hardware board:

- NeuroBayes on FPGA
- Field Programmable Gate Array: (XILINX Virtex6 VLX75T)
- Clock frequency: 250 MHz
- Approx. 1 decision per clock cycle (fully pipelined architecture)
- Probability decision output possible

NeuroBayes from Science to Industry

Predictive Analytics in High Energy Physics



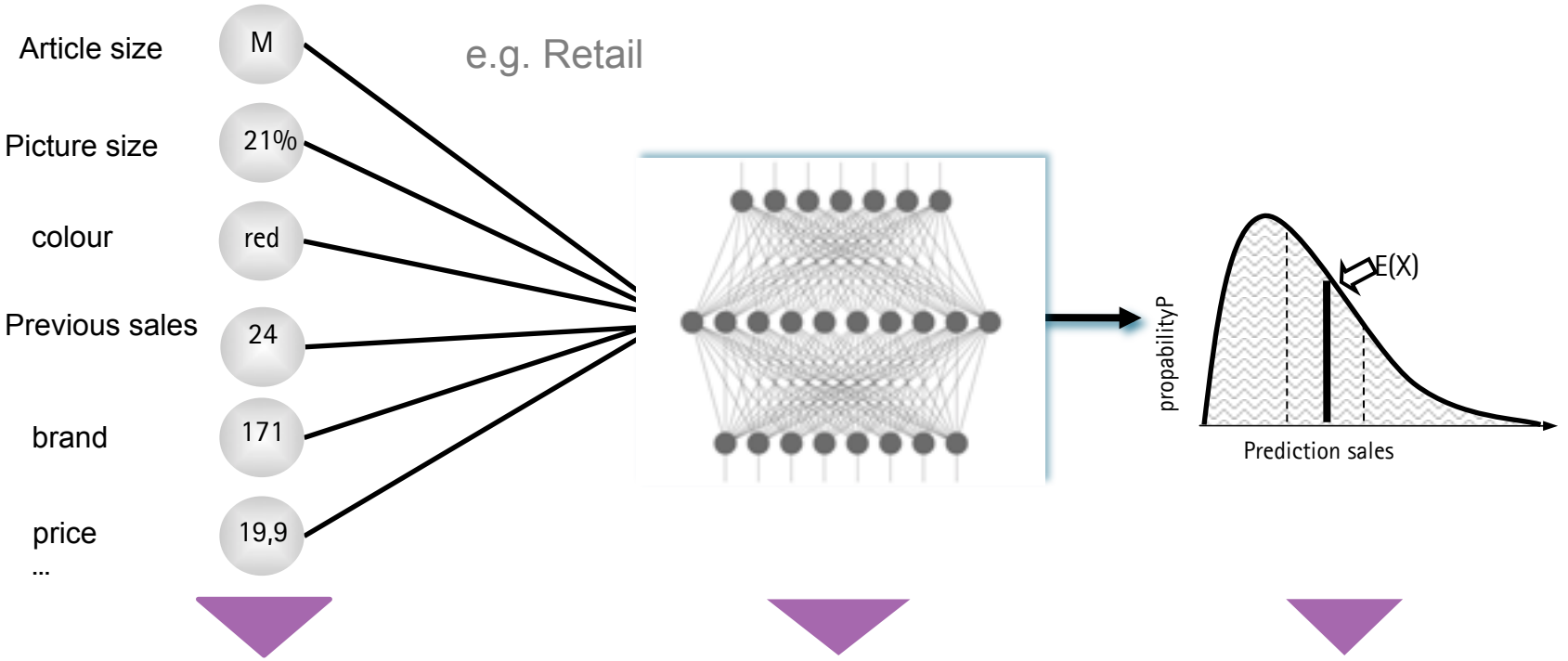
Use all available and relevant information as input, e.g. measurements from the various sub-detectors, …

NeuroBayes will extract statistically significant patterns in the data to derive the prediction.

Prediction will return the best estimator for a measurement including a statistically sound estimation of the expected spread.

NeuroBayes from Science to Industry

Predictive Analytics in industry



Use all available and relevant information as input, e.g. article properties, previous sales, etc

NeuroBayes will extract statistically significant patterns in the data to derive the prediction.

Prediction will return e.g. the most probable sales rate including a statistically sound estimation of the expected spread.

NeuroBayes allows data-driven analysis and forecasts – both in science and industry

General Overview

Mission

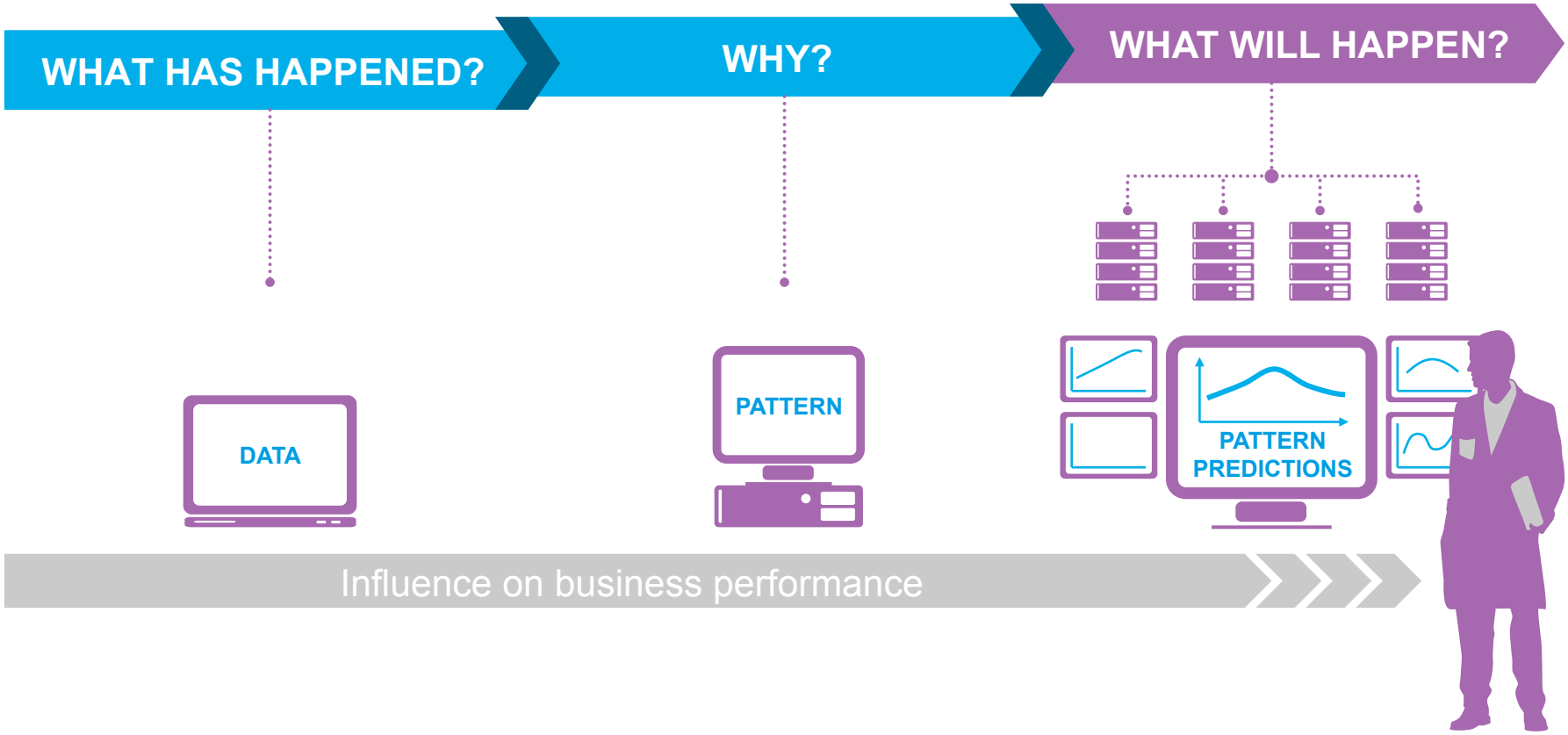
- Bring top scientific methods and thinking into business
- Forecasts and decisions purely data driven and with sound scientific methods
- Replace gut feeling (subjective priors) by objective and generalizable calculations

- Tasks are neither simple nor uninteresting. Real life is a complex system!
- Projects are very demanding! Direct comparison, fast direct feedback.

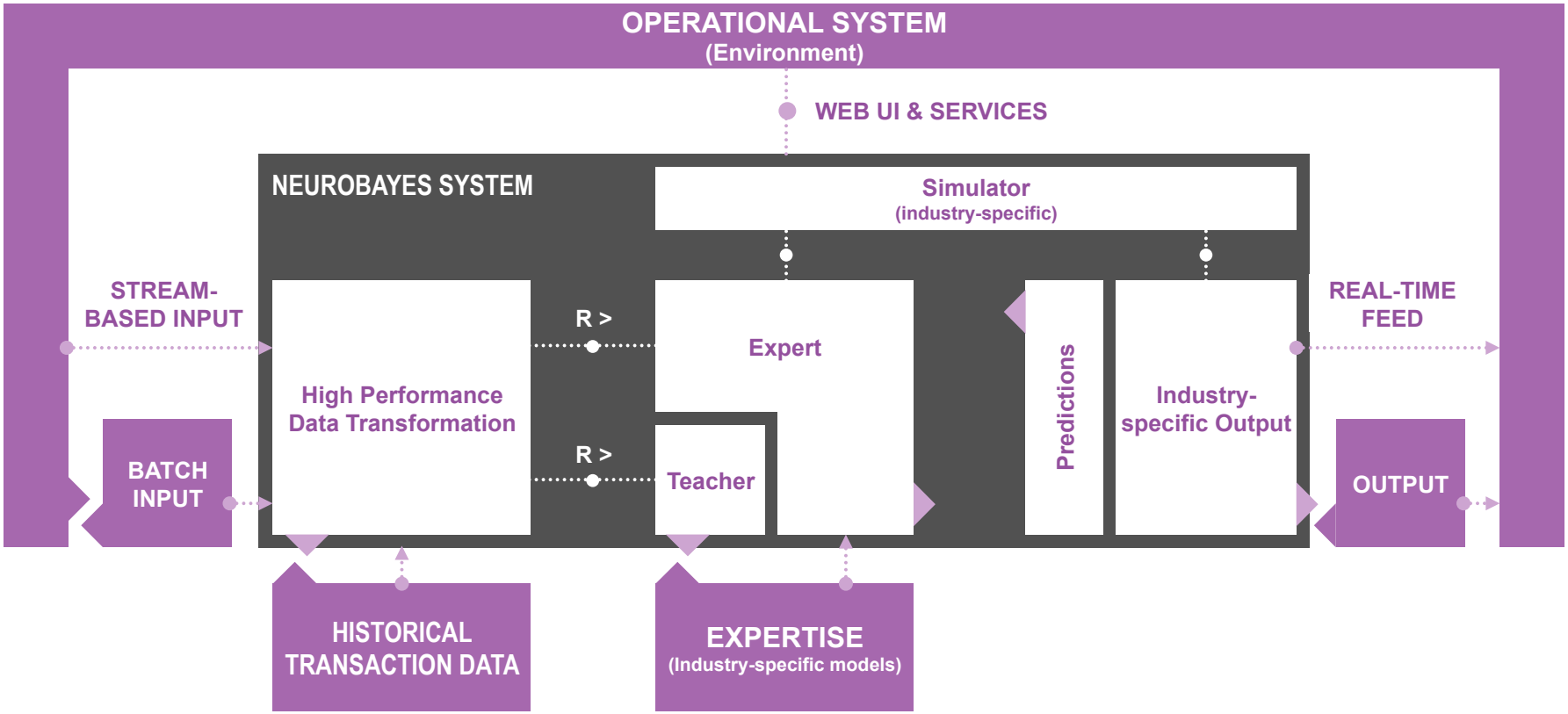
- Spirit of Blue Yonder similar to CERN.
- More than 70 PhD data scientists and 20 software engineers, mostly physicists currently dominate the total staff (currently 110)
- Very slim, extremely efficient management, sales, marketing and administration
- Largest private data science group in Europe
- (Compariable to IBM Watson core team, also 70 scientists)

Blue Yonder provides predictions --- based on data (scientifically sound – with quantified uncertainty -- testable and falsifiable, as predictive as possible)

Data Mining and conventional Business Intelligence | **Predictive Analytics**



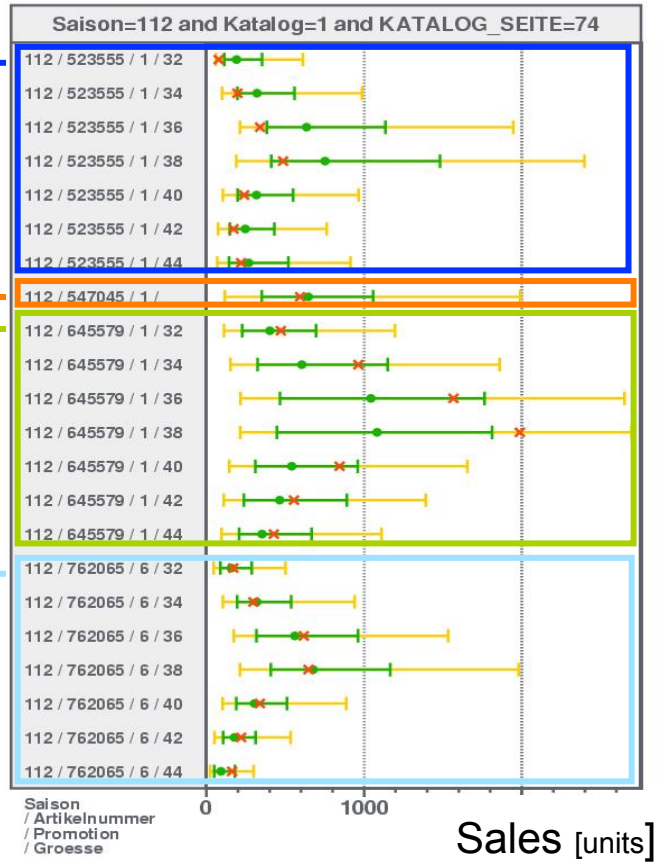
Technology Overview



- 7/24 operation needed - high safety
- Software as a Service
- Proprietary (big) data platform for development and operation

Sales Forecast Fashion

Example: OTTO Group



Per item:

- Sales forecast
- Two estimates on spread (68% and 95% confidence intervals)

Perishable goods in Supermarkets

Meat, fruit & veg, bread, diary, ….

Around 7% of all perishable foods (e.g. meat, fruit & veg., etc) have to be disposed of in German supermarkets.
That‘s about 89M tons of food wasted per year…

Insurance

e.g.
Individual risk predictions
for car insurances:

Accident probability
Claims distribution
Large claim prediction
Contract cancellation prediction

FINANCIAL TIMES DEUTSCHLAND

FORSCHEN UND

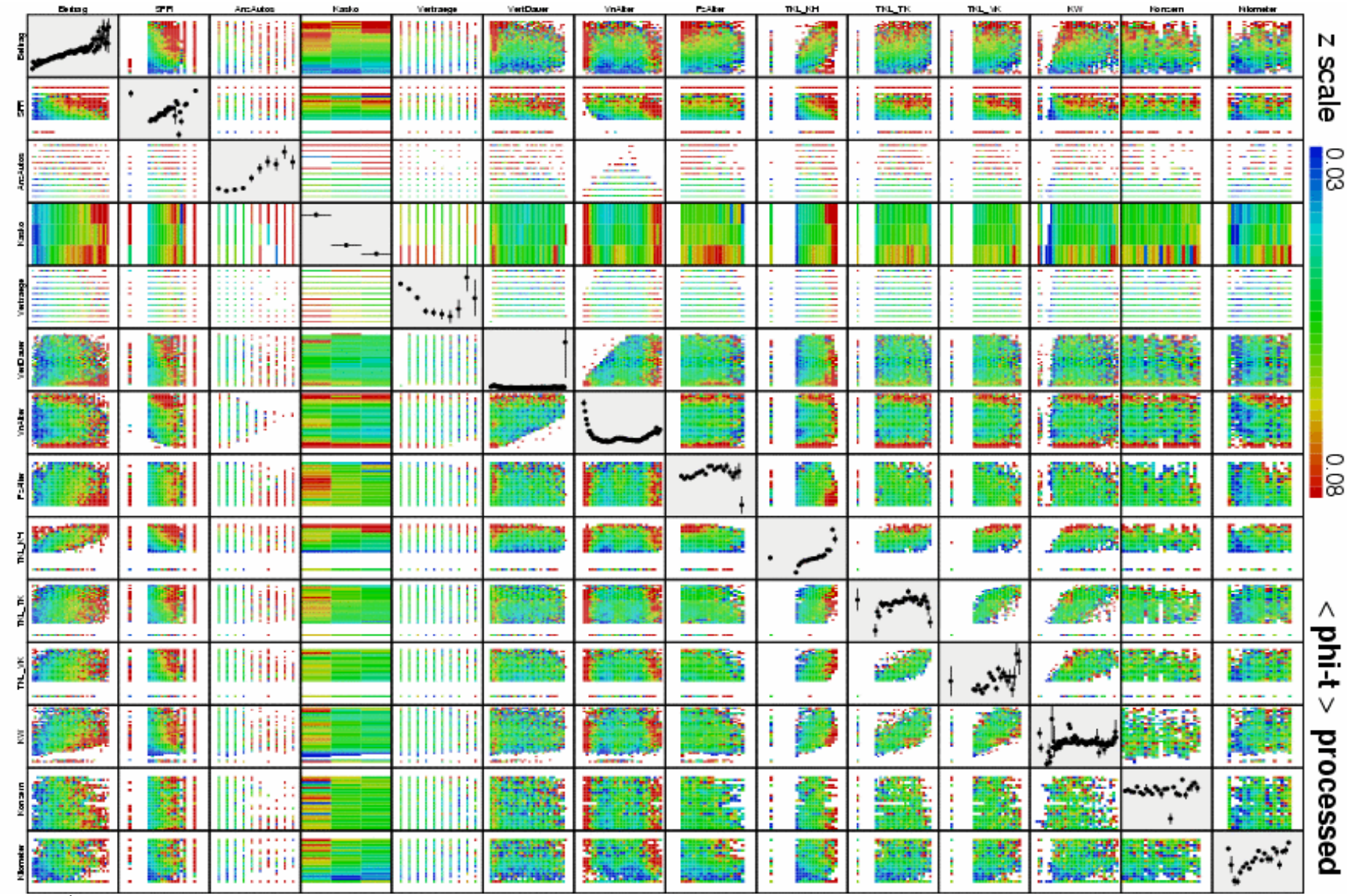
Prophetin des Knalls

Die Software „NeuroBayes" ist lernfähig – und sagt damit genauer als bisher Risiken für Kfz-Versicherungen voraus

➔ Successfully implemented at

Badisch gut versichert. BGV

Correlations to target variable „Ramler II-Plot“



NeuroBayes® delivers precise prognoses for the customer-individual number and height of claims

Premium differentiation:

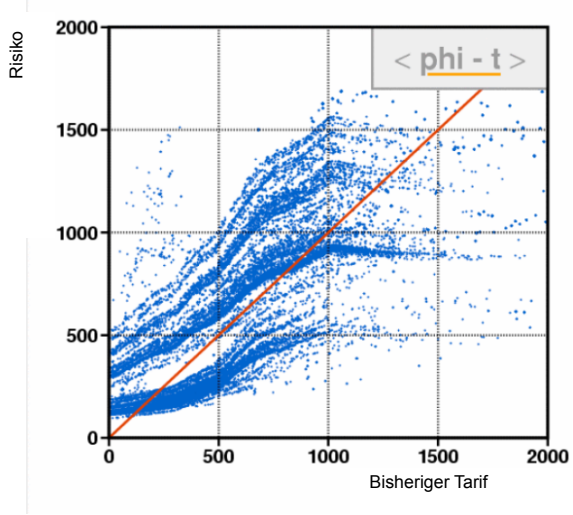
NeuroBayes® adjusts premium to customer-individual risk

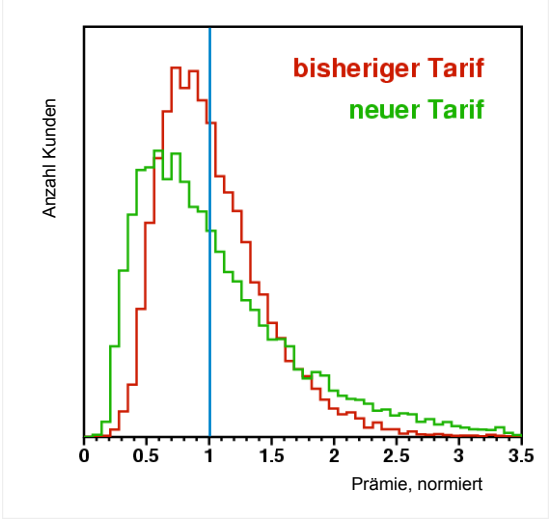
Customer structure optimisation

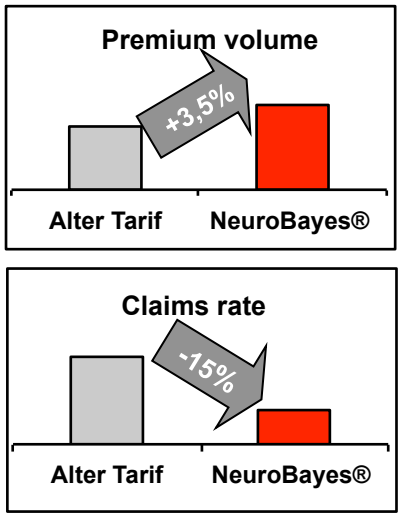
Bind your “good“ customers and take the „bad“ customers

Rentability improvement:

Simultaneously increase your total premium volume and decrease your claims rate with a more just tariff system



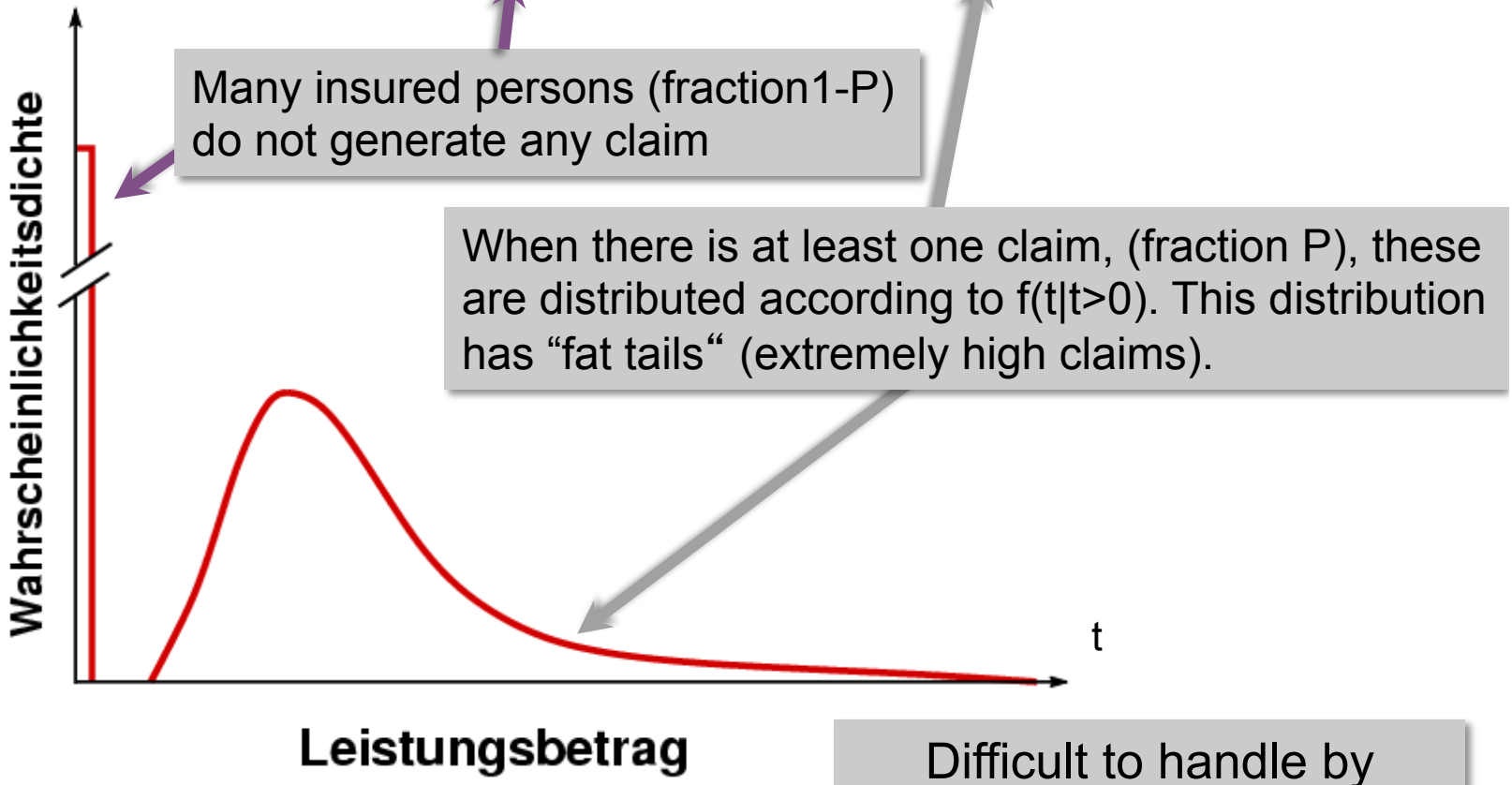




Private health insurance claims per year anything but normally distributed...

NeuroBayes® has the solution for difficult distributions of type

$$f(t) = (1 - P)\cdot \delta(t) + P\cdot\ f(t \mid t > 0)$$

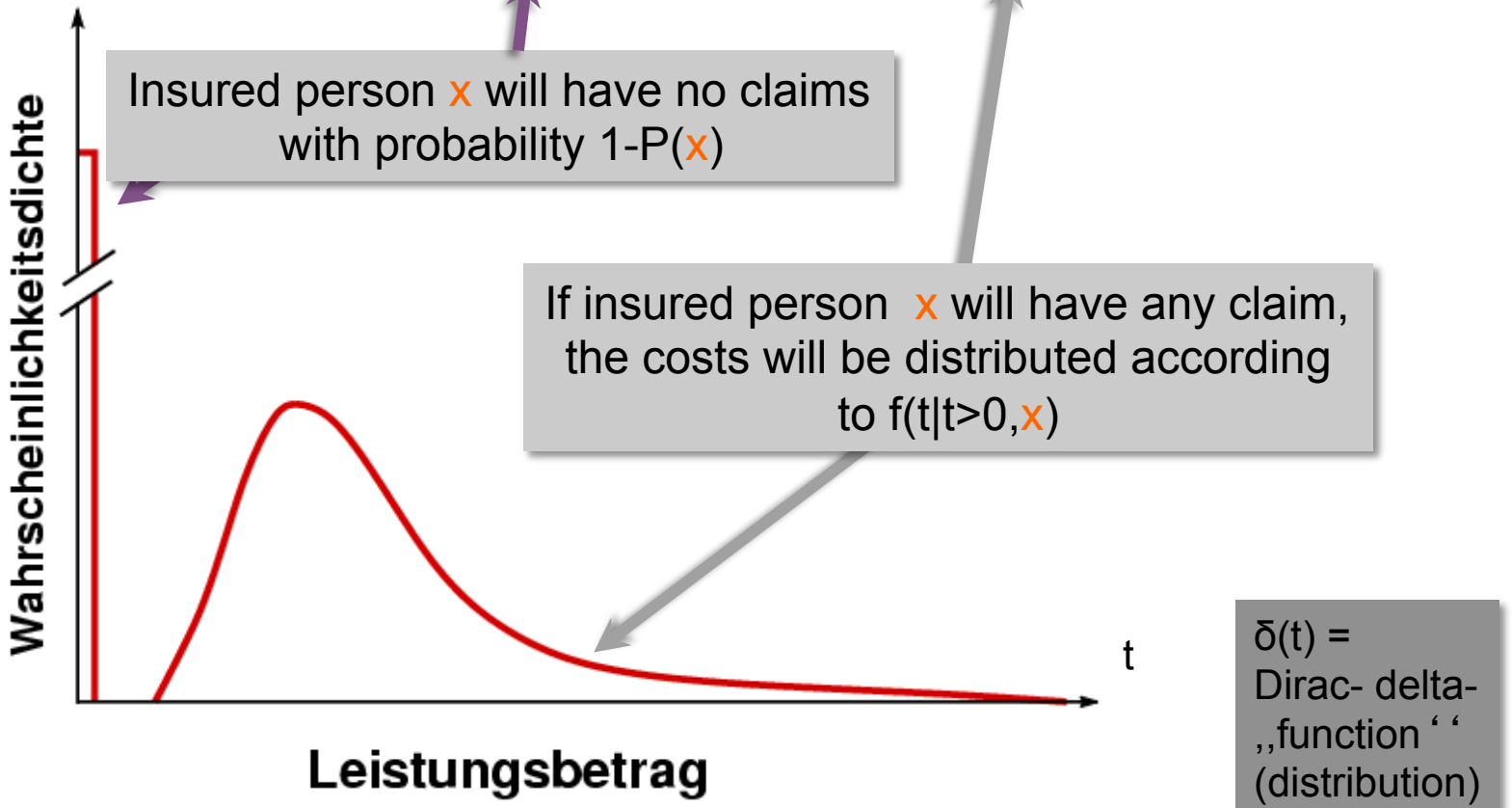


Difficult to handle by classical methods

NeuroBayes® calculates for each insured person x the individualised Bayesian probability density.

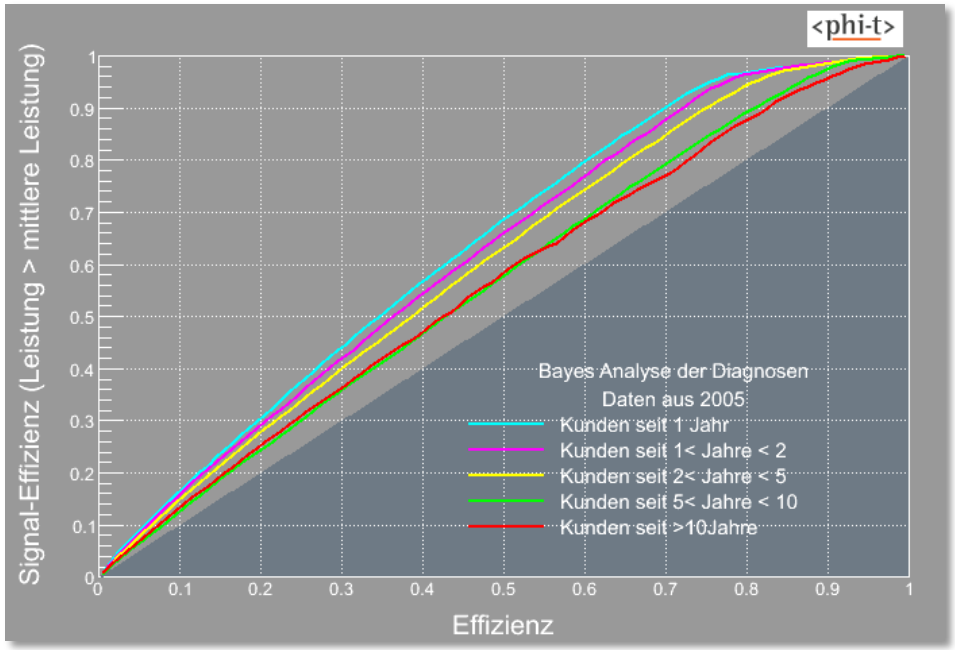
NeuroBayes® has the solution for difficult distributions of type

$$f(t \mid x) = (1 - P(x)) \cdot \delta(t) + P(x) \cdot f(t \mid t > 0, x)$$

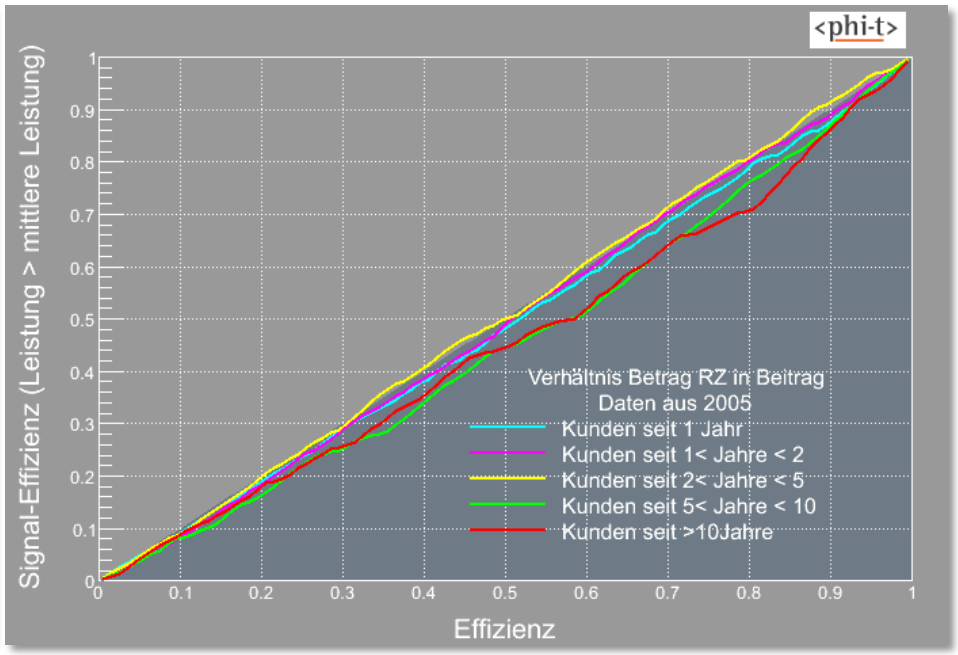


Healthcare insurance – long term prediction from anamnesis

NeuroBayes®



Expert Estimation (risk premium loading)



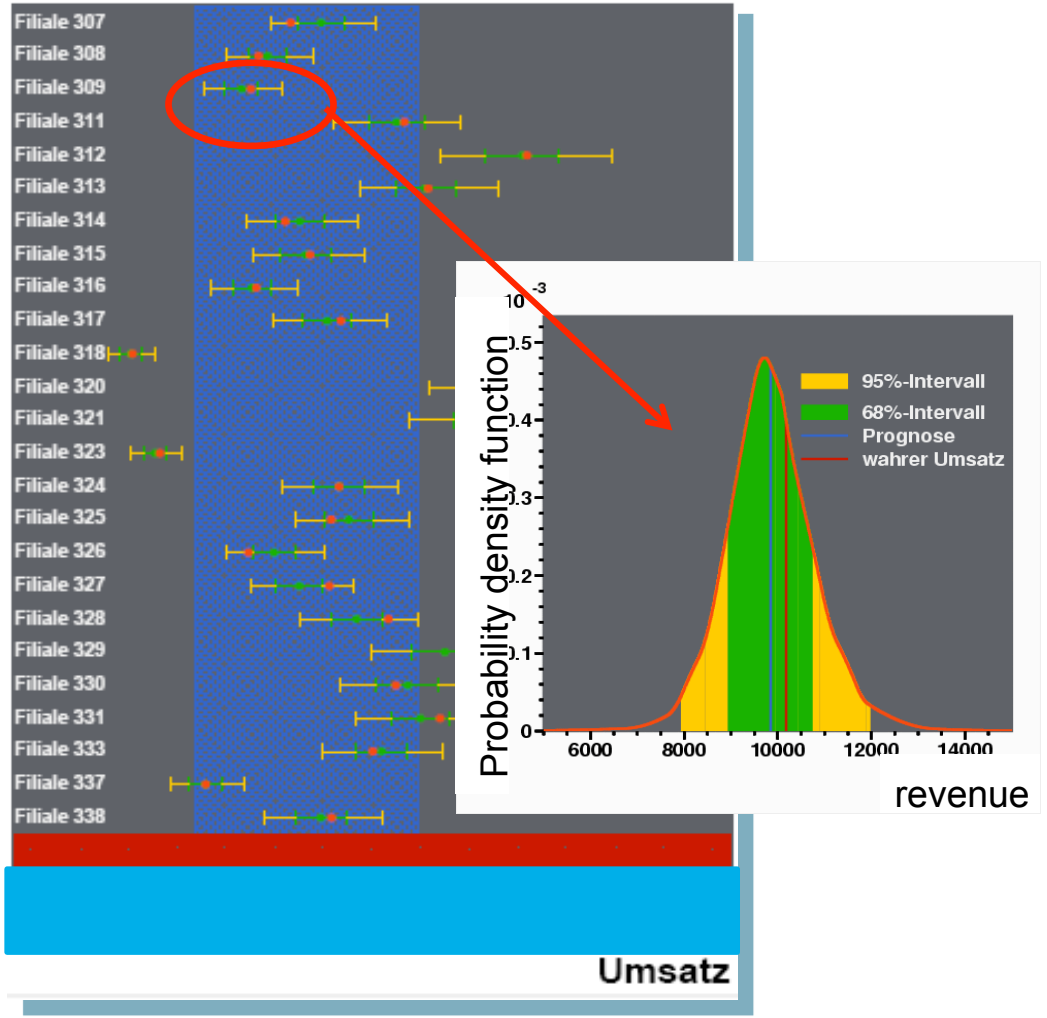
- Expert estimations are at best random –
  for patients with a long history even systematically wrong.
- NeuroBayes® forecasts costs correctly and significantly beats expert estimations more than 10 years into the future

Revenue Forecast

Example: dm– Large German drug-store chain

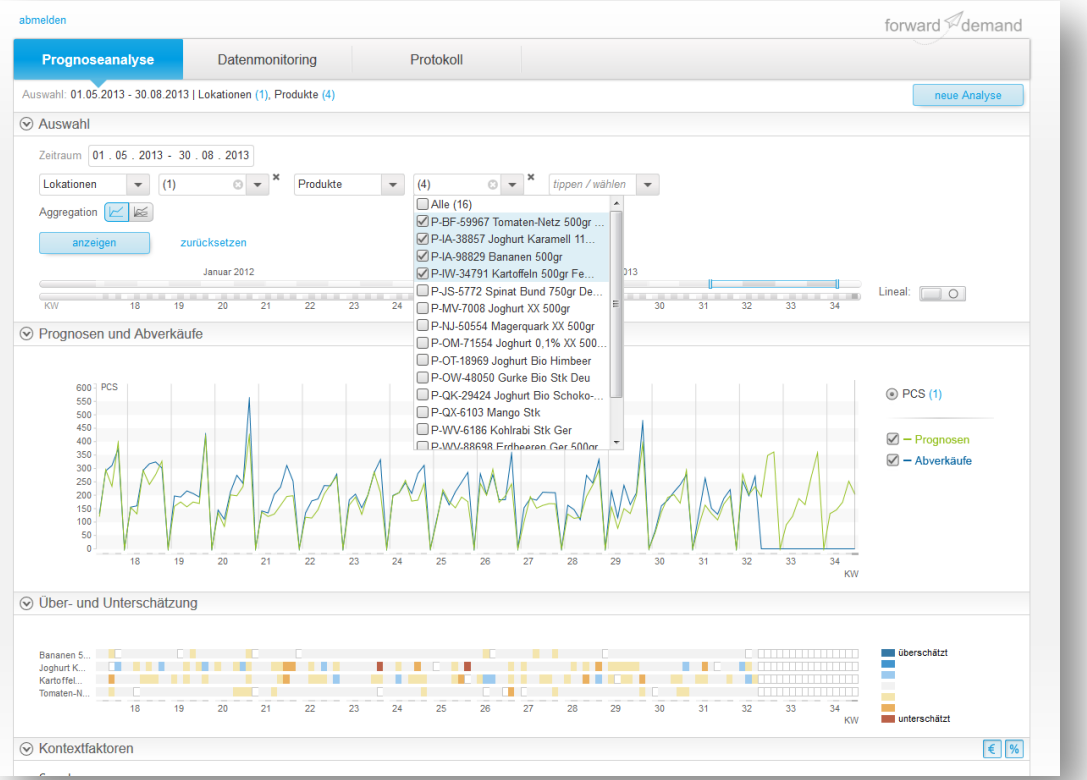
Forecasts for individual stores

- Prediction of the full probability density function.
- Precise forecast of the expected revenue including exptected spread
  (68% and 95% confidence intervals)



forward demand

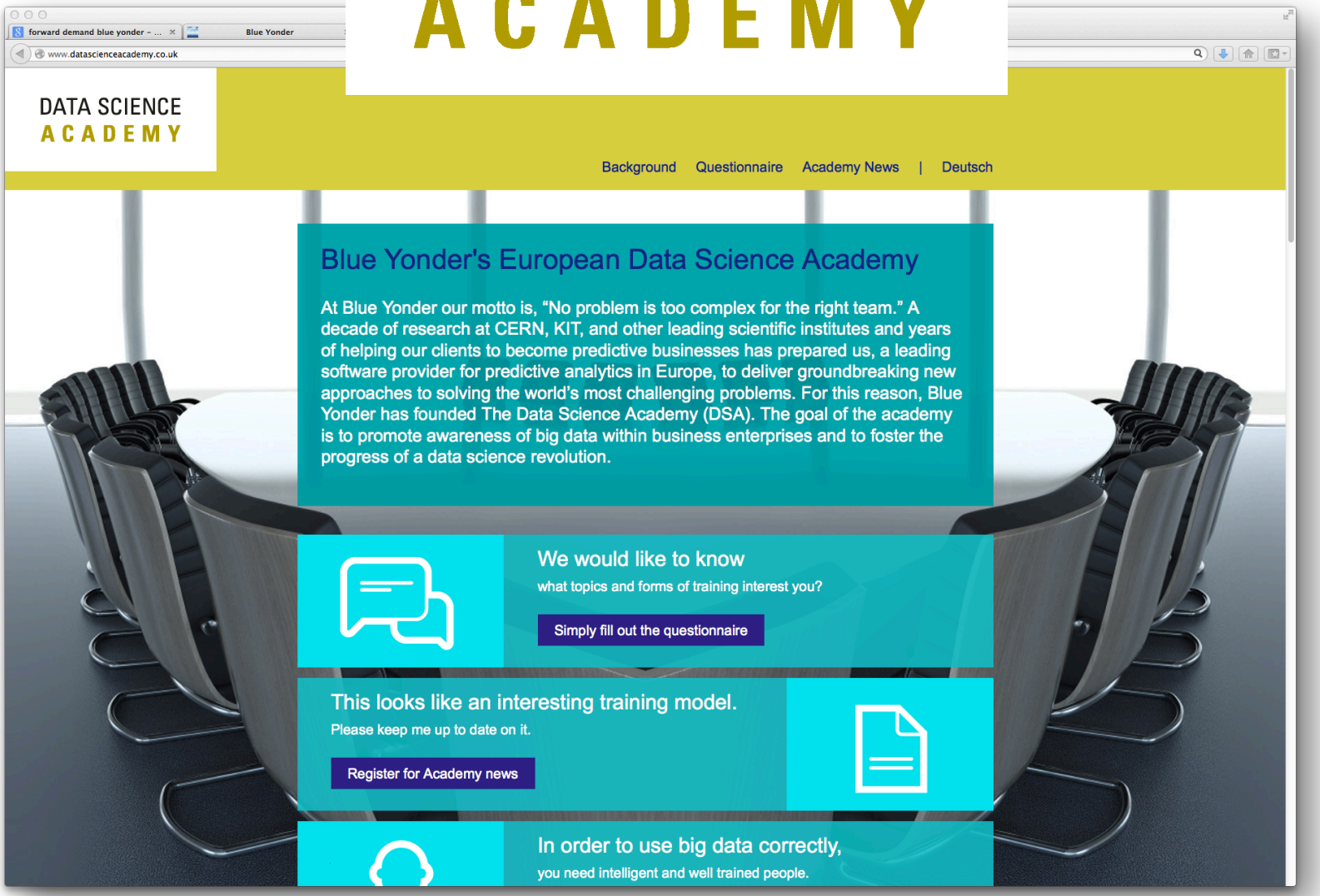
- Just released: First big data predictive analytics standard software solution
- Precise sales forecasts for retail and CPG
- Access to big data predictive analytics for B2B end users
- Easy handling through intuitive web-UI
- Software-as-a-Service allows usage of Forward Demand without high in-advance investments into software, infrastructure or highly skilled personell



www.blue-yonder.com/forwarddemand

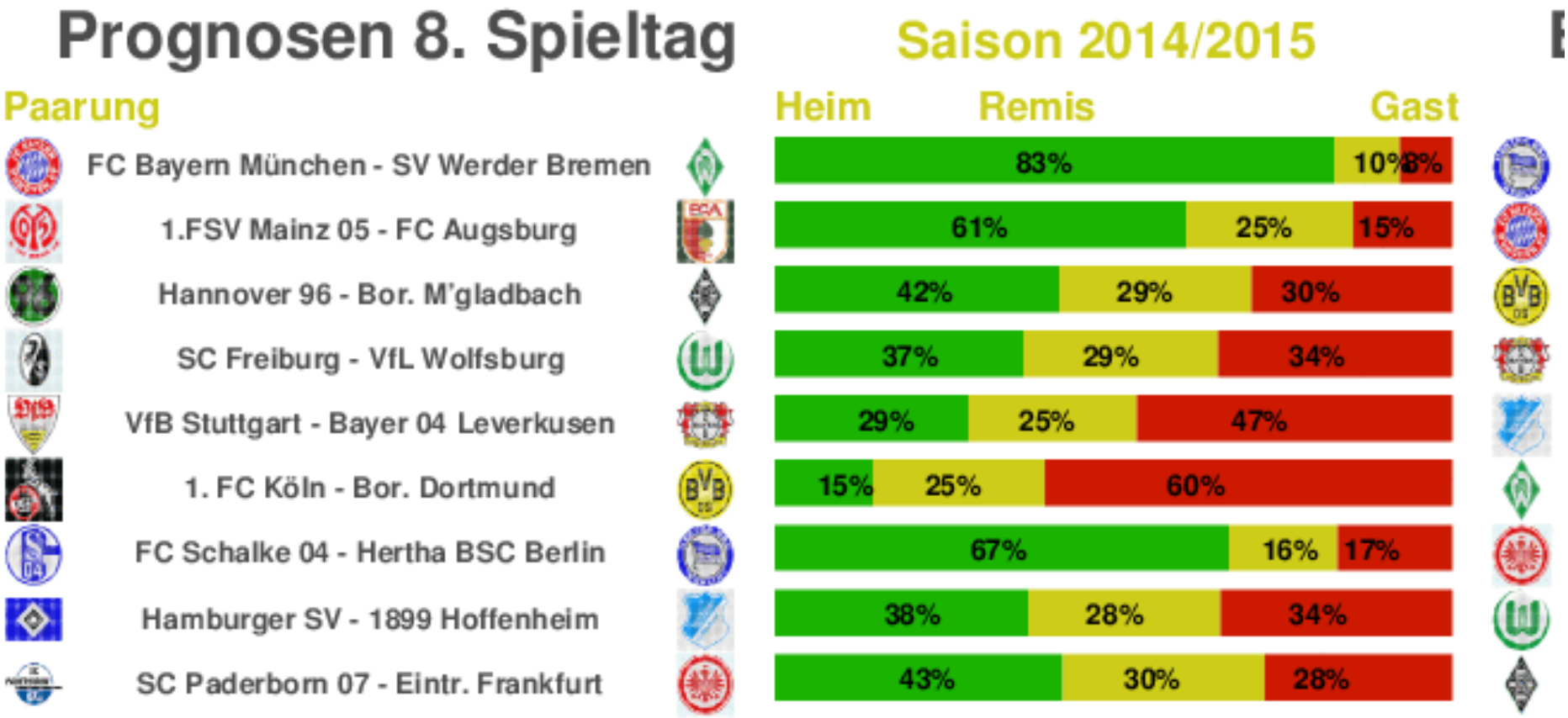
Just launched:

DATA SCIENCE ACADEMY



www.datascienceacademy.eu

Prognosis of sports events from historical data: NeuroNetz*er*



Results: Probabilities for home - tie - guest

Blue Yonder trends in data analytics: data bases

- Excel is standard in many departments in most companies
- Graphical data representations are largely unavailable / unknown
- SQL already is advanced for many users
- Many data warehouses are “write-only“, analysis access too slow / restricted

The advanced stuff is:

- Vertical data bases! (columnar ntuples for 25 years already in HEP)
- In-memory data bases (mostly SQL-based)
- Not one optimal solution for all problems of the world.
- MAP/REDUCE (Hadoop)

Blue Yonder trends in data analytics...

- Fast interactive parallel data analysis / model development important for professional data scientists:

- PIAF/PROOF too much forgotten
- MAP/REDUCE not the answer to everything (good: CPU at data)

- Combine goodies of both...

- root / C++ too complicated and not flexible enough (e.g. store new columns)

- Simplicity (important even for super data scientists): C++ → Python
- Python: numpy, pandas etc very effective, CPU time is (largely) NOT a limitation

- Machine Learning community often thinks too deterministic, no good understanding of (no interface for) statistical errrors and weights

Blue Yonder trends in data analytics...

Article **Why cutting edge technology matters for Blue Yonder solutions**
► http://www.blue-yonder.com/en/resource-center/research-papers.html
gives overview on our view on algorithms.

We consider as important keywords:

To know what makes a good prediction…
Domain knowledge in different industries
NeuroBayes
Neural Networks
Bayesian statistics
Deep learning
Reinforcement learning
Correlation and Causality -- automatic interventions
Big data
Parallelization -- enhance NeuroBayes to arbitrary large training sets

Blue Yonder – fastest growing BI company in Germany

Aufsteiger: Top 10-Anbieter von Business-Intelligence-Software nach Umsatzzuwachs

Unternehmen	Umsatzwachstum 2011-2012	Softwareumsatz 2012 (in Mio. €)
Blue Yonder	175%	3,3
Talend	100%	8,5
Splunk	100%	7,0
Rapid-I	78%	1,6
Jedox	43%	3,0
Tibco Spotfire	40%	18
LucaNet	39%	6,4
Bissantz	38%	6,2
SAP	37%	240
Datawatch	35%	2,3
BI-Gesamtmarkt	**13,0%**	**1190**

Quelle: „Der Markt für Business Intelligence in Deutschland 2012", BARC.

BARC

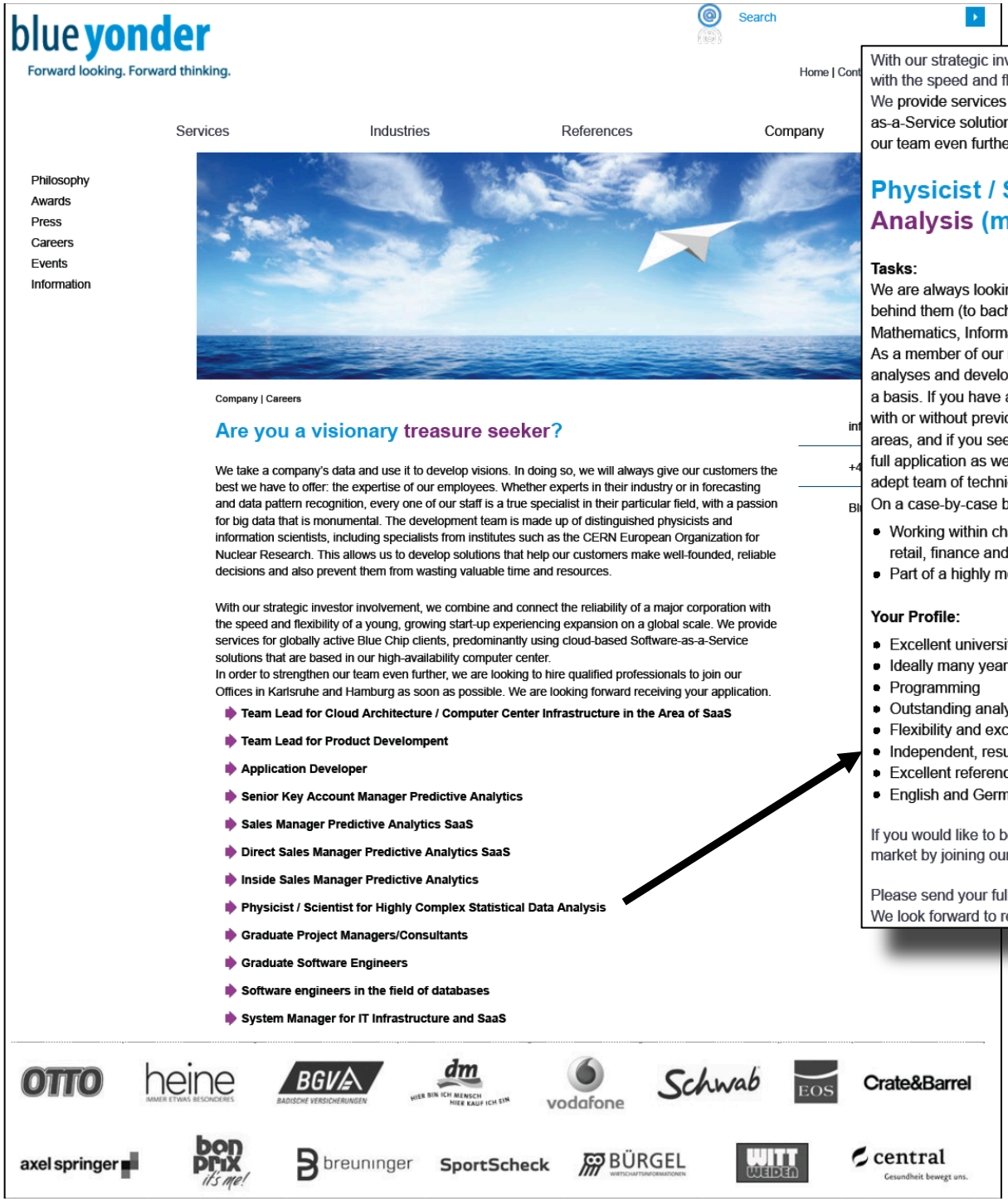
The Business Application Research Center (BARC) found that Blue Yonder is the fastest growing BI software company in Germany.

With 175% turnaround increase in 2012 Blue Yonder is leading the field compared to 250 competitors in the area of Business Intelligence and data management.

Career opportunities in Karlsruhe, Hamburg and London:

http://www.blue-yonder.com/unternehmen/karriere.html

http://www.blue-yonder.com/en/company/career.html



With our strategic investor involvement, we combine and connect the reliability of a major corporation with the speed and flexibility of a young, growing start-up experiencing expansion on a global scale. We provide services for globally active Blue Chip clients, predominantly using cloud-based Software-as-a-Service solutions that are based in our high-availability computer center. In order to strengthen our team even further, we are looking to hire qualified professionals in the area of data analysis.

Physicist / Scientist for Highly Complex Statistical Data Analysis (m/w)

Tasks:

We are always looking to hire motivated individuals who have an outstanding university education behind them (to bachelor degree, master's degree or doctorate level) in the subjects of Physics, Mathematics, Information Sciences, Business Informatics or similar.
As a member of our rapidly growing company, you will be responsible for creating highly complex data analyses and developing forecasting models for our clients using the very latest statistical methods as a basis. If you have a strong interest in data analysis, multivariate statistics and programming (either with or without previous professional experience), ideally coupled with some expertise within these areas, and if you see yourself becoming part of an ambitious high-tech company, please send us your full application as we want to hear from you. You will be part of a young, intelligent and outstandingly adept team of technical scientists based in Karlsruhe or Hamburg.
On a case-by-case basis, we also offer the chance to receive financial support for your doctoral thesis.

- Working within challenging and demanding data analysis projects for various industries such as retail, finance and insurance
- Part of a highly motivated team of data analysts and software developers

Your Profile:

- Excellent university degree
- Ideally many years of experience involving data analysis
- Programming
- Outstanding analytical skills
- Flexibility and excellent team spirit
- Independent, results-oriented approach with a will to succeed
- Excellent references
- English and German language skills, additional languages advantageous

If you would like to become part of a demanding, highly-innovative and extremely sustainable software market by joining our young and highly-motivated team, we want to hear from you!

Please send your full set of application documents in electronic form to **jobs@blue-yonder.com**.
We look forward to receiving your application.

www.blue-yonder.com

Blue Yonder: Awards for Big Science Startup

DLD 2013: Best Enterprise Solution

DIP 2014 in category large enterprises: OTTO for employing Blue Yonder technology in the complete product lifecycle

Der Deutsche Innovationspreis

Top Retail Product 2011/2012

Retail Technology Award 2012

3 time winner of the Data Mining Cup

bwcon Hightech Award 2012

Finalist 2012
Finalist 2013

Special Prize Deutsche Boerse 2012