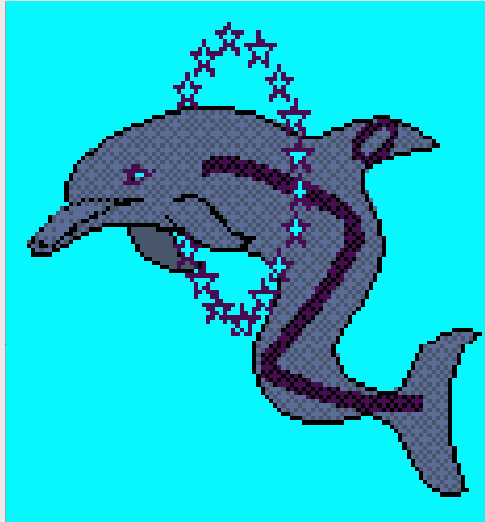




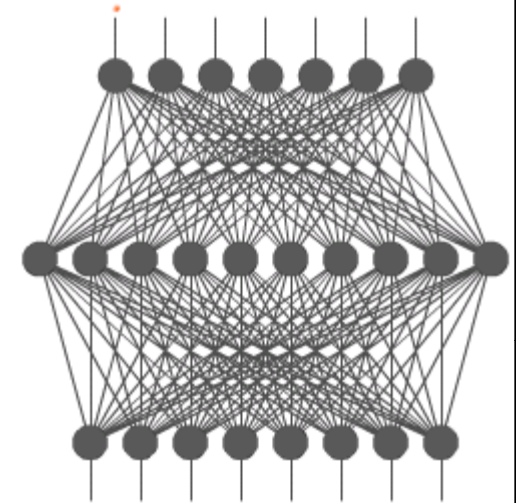
Universität Karlsruhe

CETA



From Delphi...

...to | <phi-t><sup>®</sup>  
Physics Information Technologies



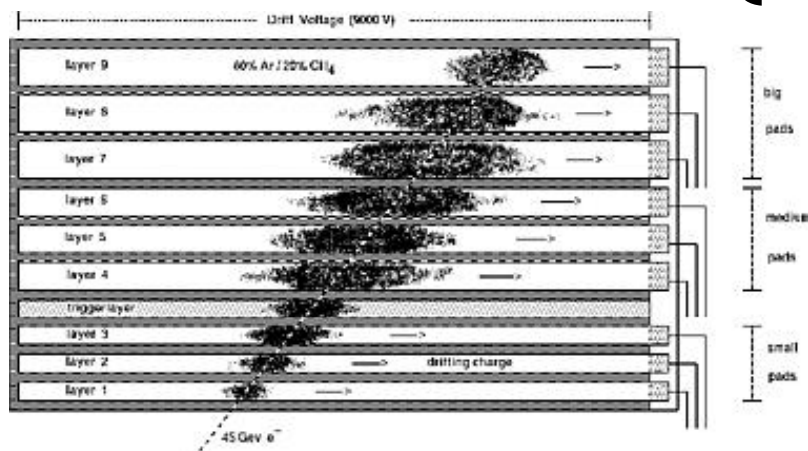
## Spin-Off from Physics Research to Business

Prof. Dr. Michael Feindt  
KCETA - Centrum für Elementarteilchen- und Astroteilchenphysik  
IEKP, Universität Karlsruhe, Karlsruhe Institute of Technology KIT  
& Phi-T GmbH, Karlsruhe

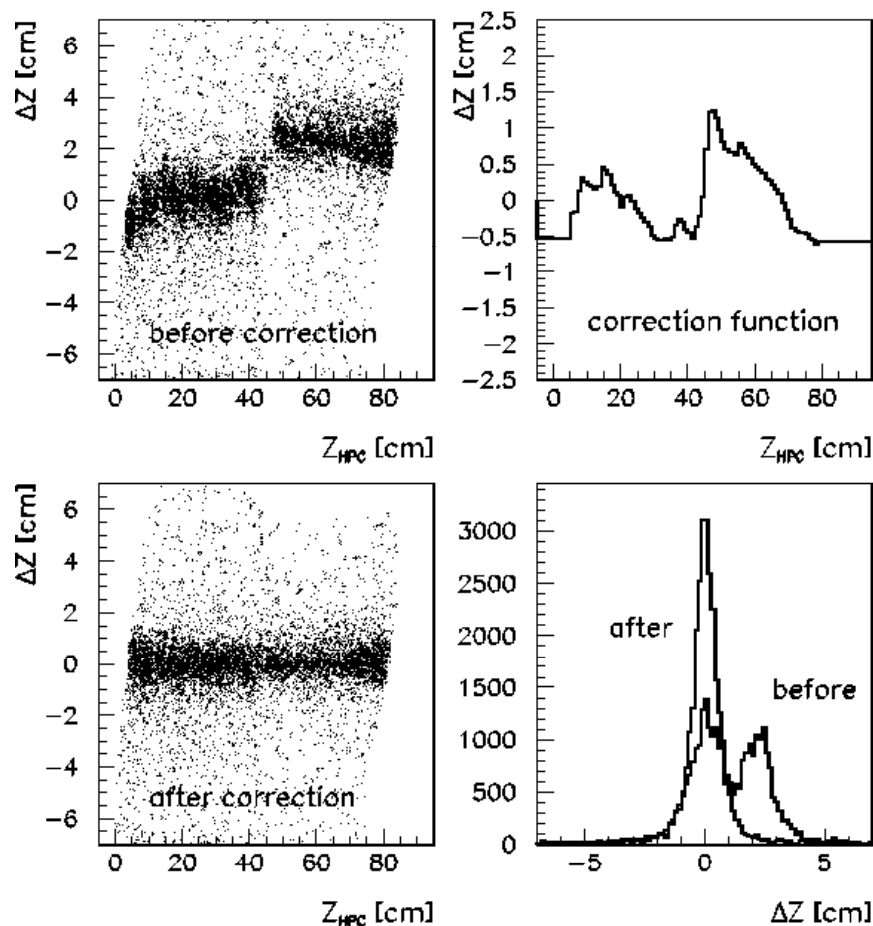
DELPHI Closing Meeting

CERN, May 29, 2008

# HPC (1991-92)



HPC Z – Resolution for Module 111



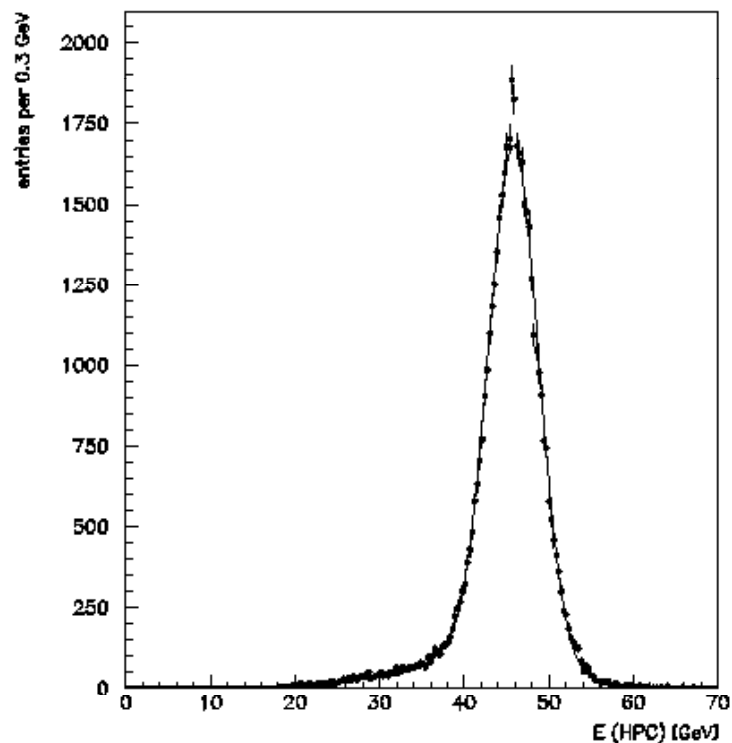
Initially energy and spatial resolutions catastrophic

Solve problems by looking **IN DETAIL** (144 modules) into the real **DATA**.  
Correct by software

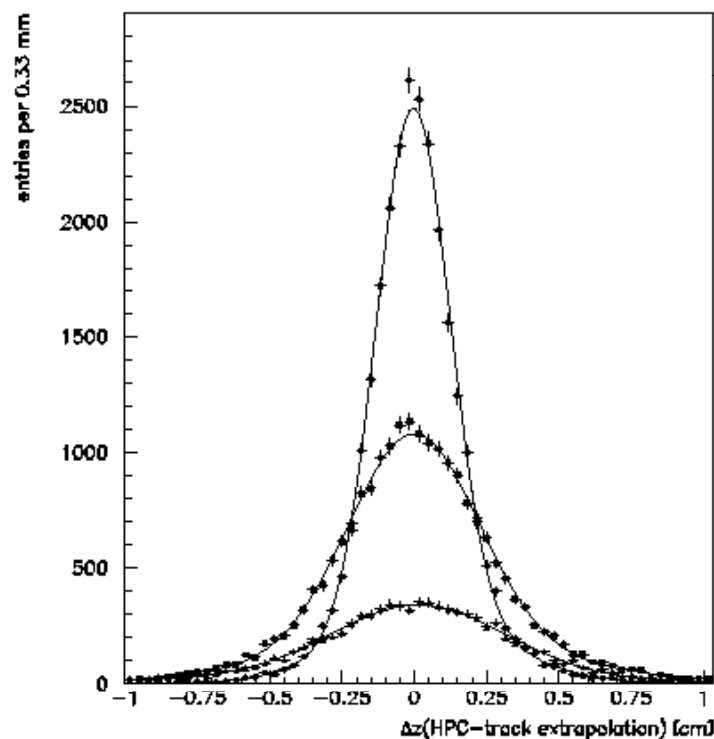
# HPC

→ design performance almost reached, physics analysis possible

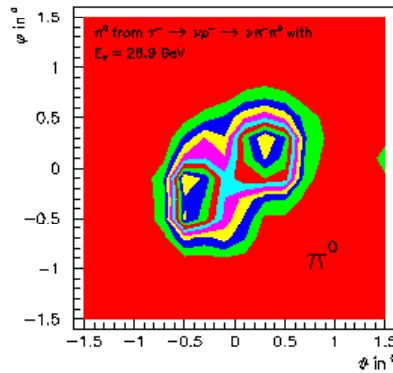
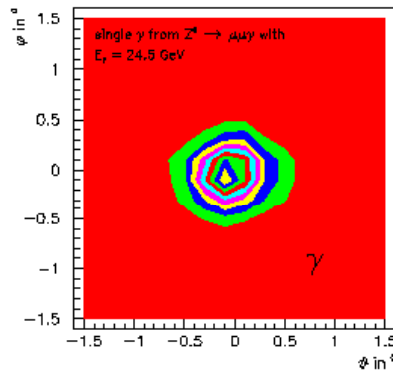
HPC energy resolution



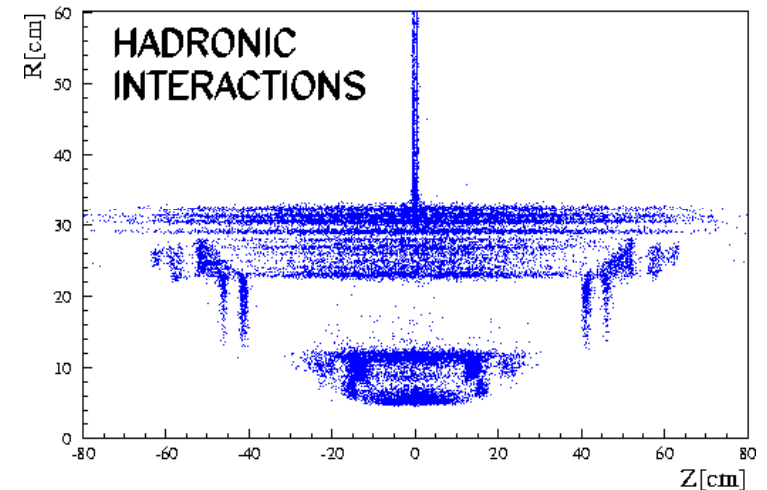
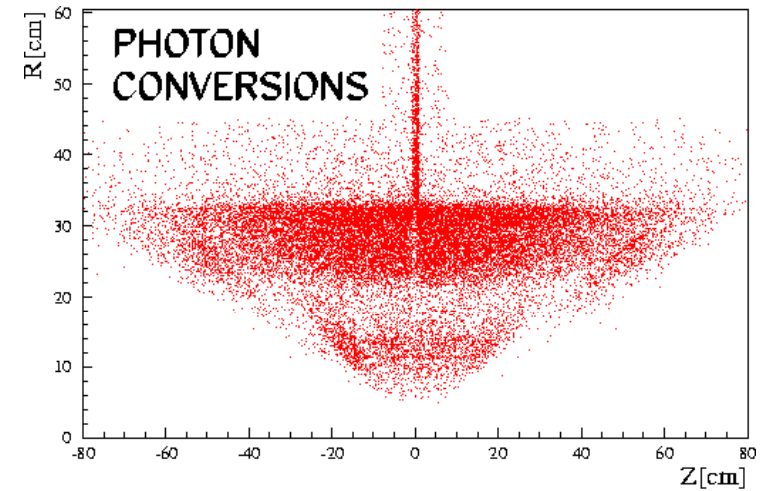
HPC z-resolution



# Introduction of Software Tasks



e.g.  
ELEPHANT,  
MAMMOTH  
packages



# (Inclusively reconstr.) B physics

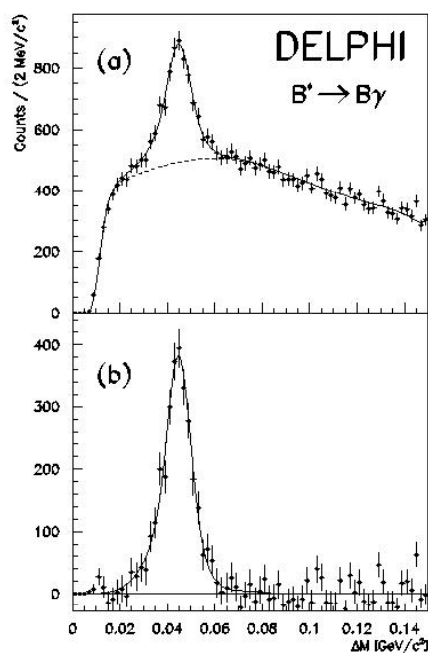


Figure 3:  $H\gamma - H$  mass difference distribution. (a) The data are represented by points, the line corresponds to the fit result using background and signal shapes predicted by simulation. (b) The  $\Delta M(H^* - H)$  signal after background subtraction. The curve shows the expectation from the simulation with a value of  $45.5 \text{ MeV}/c^2$  and an isospin splitting of  $3.5 \text{ MeV}$ .

$B^*$

$B^{**}$

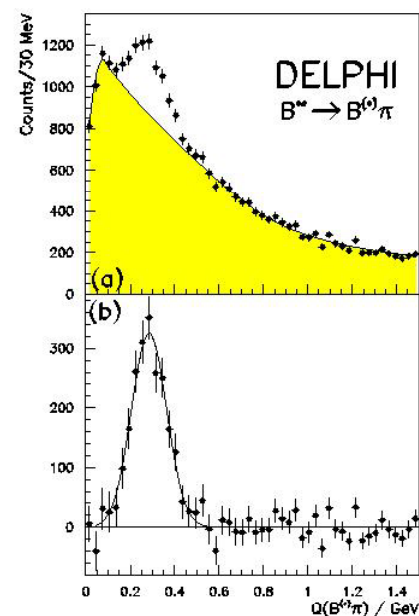
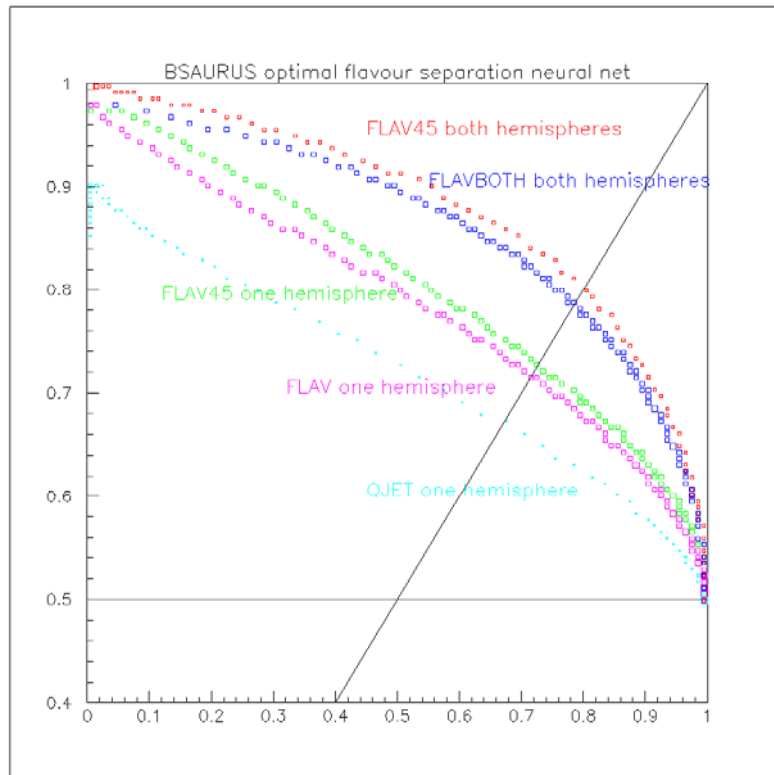
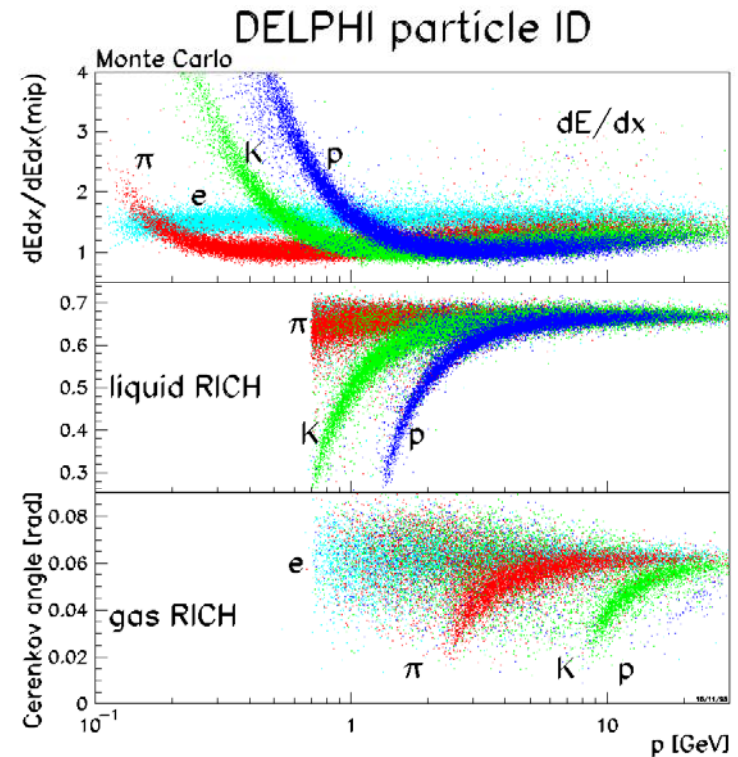


Figure 1: (a) Distribution of the Q-value of  $B^{(*)}\pi$  pairs (data points) along with the Monte Carlo expectation without  $B^{**}$  production (shaded area). Q is defined as  $Q = m(B^{(*)}\pi) - m(B^{(*)}) - m(\pi)$ . (b) Background subtracted  $B^{(*)}\pi$  pair Q-value distribution. The fit is a simple Gaussian.

# Start to work with neural networks (in 1993)



**BSAURUS: b flavour tagging**



**MACRIB: Combined Particle ID**

# Gained lots of experience with neural network applications in DELPHI (1993–2003)

Electron id (ELEPHANT)

Kaon, proton id (MACRIB)

(BSAURUS::

Optimisation of resolutions inclusive B- E,  $\pi$ ,  $\eta$ , Q-value

Inclusive B production flavour tagging

Inclusive B decay flavour tagging

Inclusive B<sup>+</sup>/B<sup>0</sup>/B<sub>s</sub>/Lambda<sub>b</sub> separation

B<sup>\*\*</sup>, B<sub>s</sub><sup>\*\*</sup> enrichment

B fragmentation function

Limit on B<sub>s</sub>-mixing

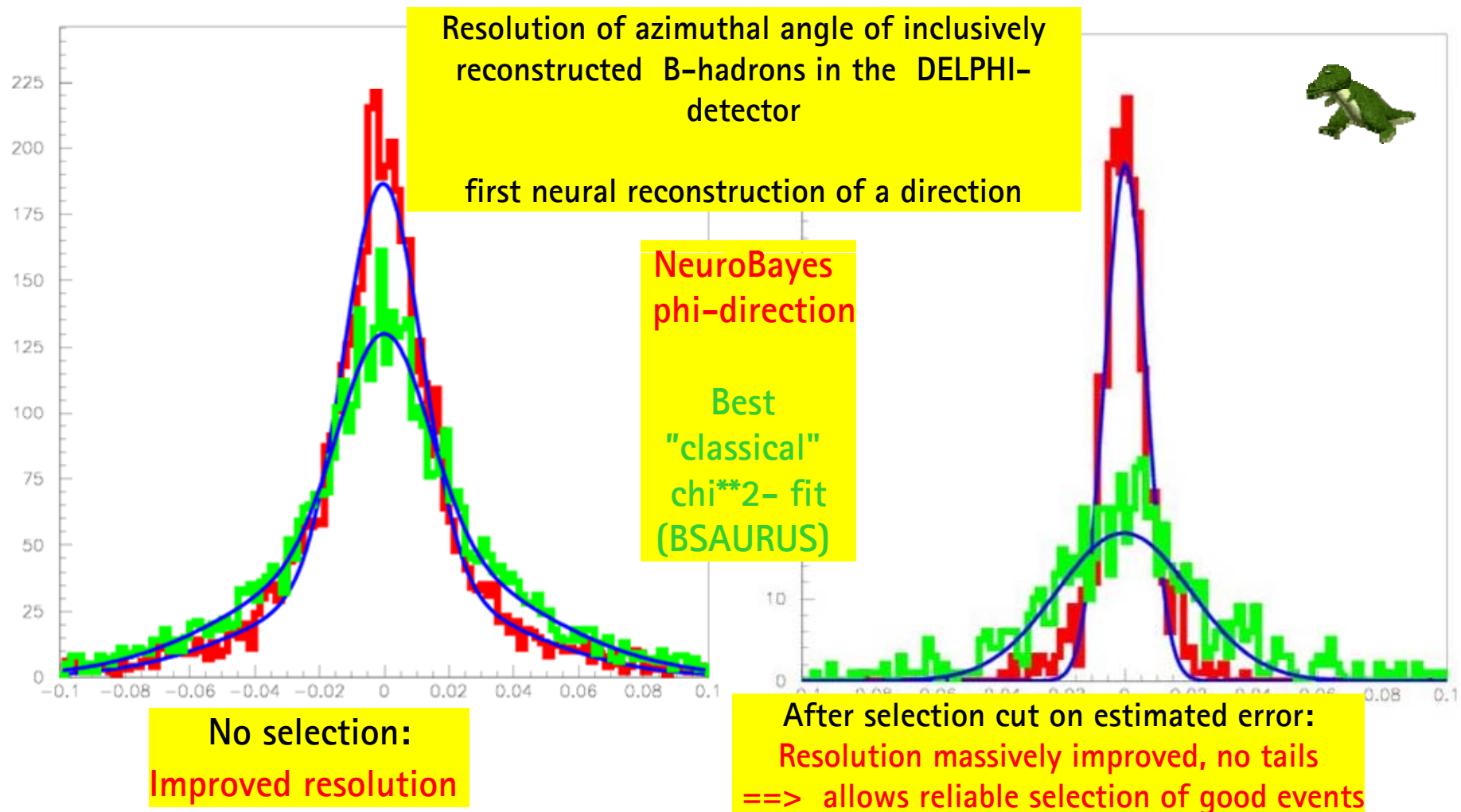
B<sup>0</sup>-mixing

B- F/B-asymmetry

B- $\rightarrow$  wrong sign charm)



# Neural reconstruction of a direction





## Naïve neural networks and criticism

We've tried that but it didn't give good results

- Stuck in local minimum
- Learning not robust

We've tried that but it was worse than our 100 person-years analytical high tech algorithm

- Selected too naive input variables
- Use your fancy algorithm as INPUT !

We've tried that but the predictions were wrong

- Overtraining: the net learned statistical fluctuations

Yeah but how can you estimate systematic errors?

- How can you with cuts when variables are correlated?
- Tests on data, data/MC agreement possible and done.

**Address all these topics and build a professional robust and flexible neural network package for physics and general prediction tasks:**

**NeuroBayes<sup>®</sup>**

**(2000–2002)**

**turned out to be a very strong tool**

After having looked outside the ivory tower, realized:

**These methods are not only applicable in physics**



**<phi-t>**: Foundation of private company out of University of Karlsruhe, initially sponsored by exist-seed-programme of the federal ministry for Education and Research BMBF

# History

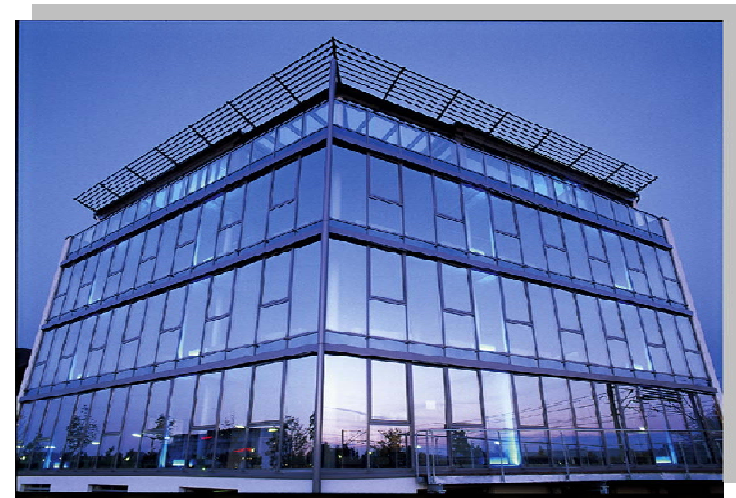
2000–2002 NeuroBayes<sup>®</sup>–specialisation  
for economy at the University  
of Karlsruhe

Oct. 2002: GmbH founded,  
first industrial projects

June 2003: Removal into new office  
199 sqm IT-Portal Karlsruhe

May 2008: Expansion to 700 sqm office  
Foundation of sub-company  
Phi-T products&services

Exclusive rights for NeuroBayes<sup>®</sup>  
Staff all physicists (almost all from HEP)  
Continuous further development of NeuroBayes<sup>®</sup>



# <phi-t><sup>®</sup> Customers:

Physics Information Technologies

(among others):

BGV and VKB car insurances  
Lupus Alpha Asset Management  
Otto Versand (mail order business)  
Thyssen Krupp (steel industry)

AXA and Central health insurances  
dm drogerie markt (drugstore chain)  
Libri (book wholesale)

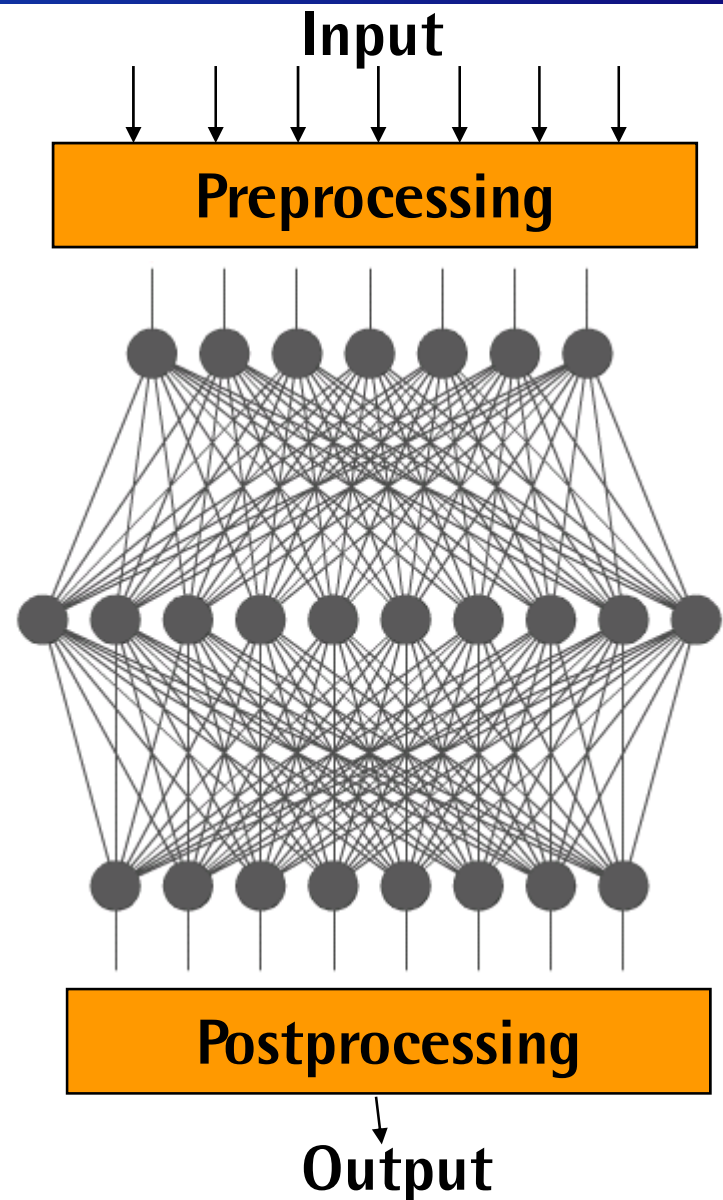


# NeuroBayes<sup>®</sup> principle

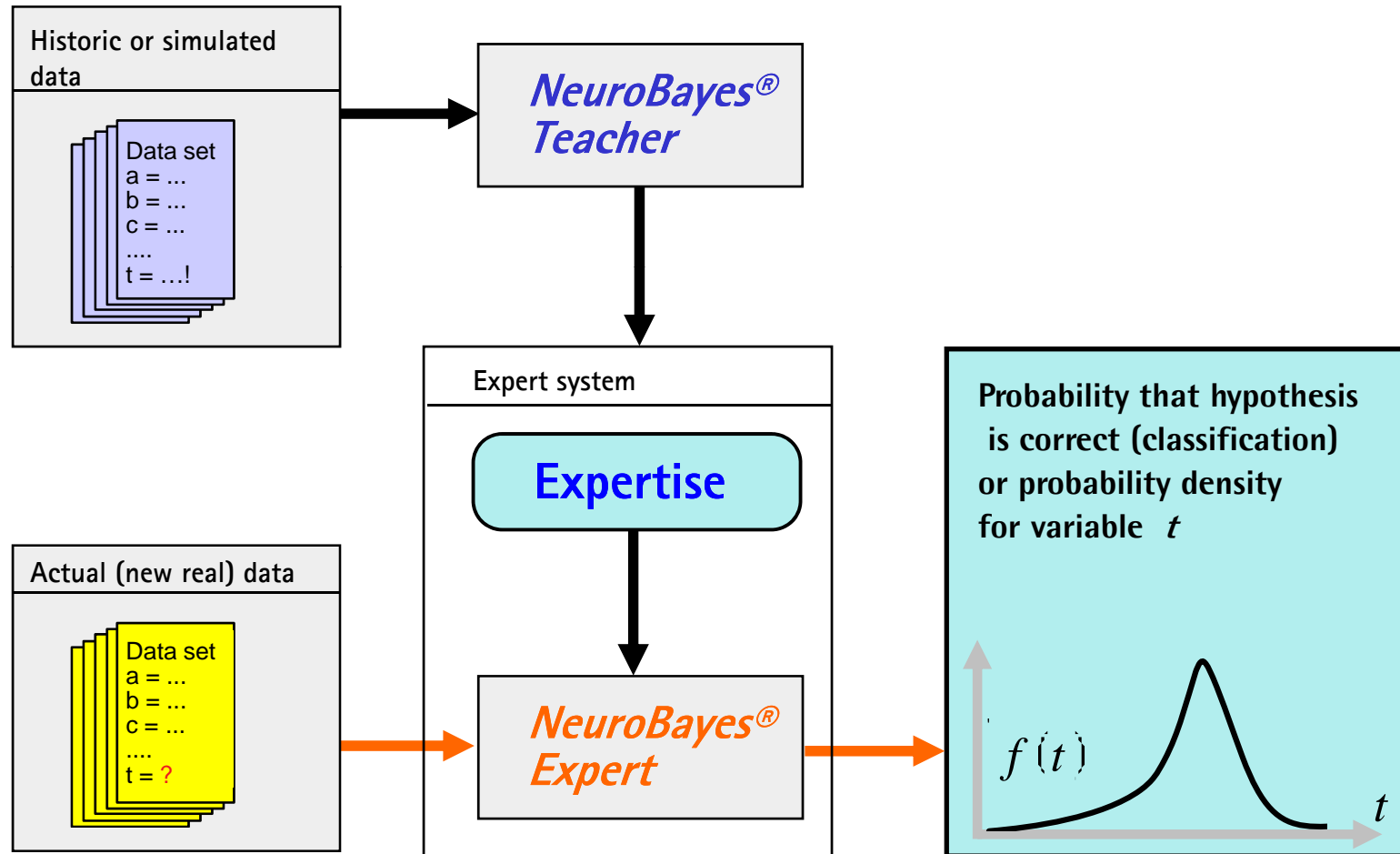
**NeuroBayes<sup>®</sup> Teacher:**  
Learning of complex relationships from existing data bases (e.g. Monte Carlo)

**NeuroBayes<sup>®</sup> Expert:**  
Prognosis for unknown data

Significance control



# How it works: training and application



# NeuroBayes<sup>®</sup> 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 event is a Higgs candidate.
- > Germany will become soccer world champion in 2010.
- > Customer Meier will have liquidity problems next year.
- > This equity price will rise.



# NeuroBayes<sup>®</sup> task 2: Conditional probability densities

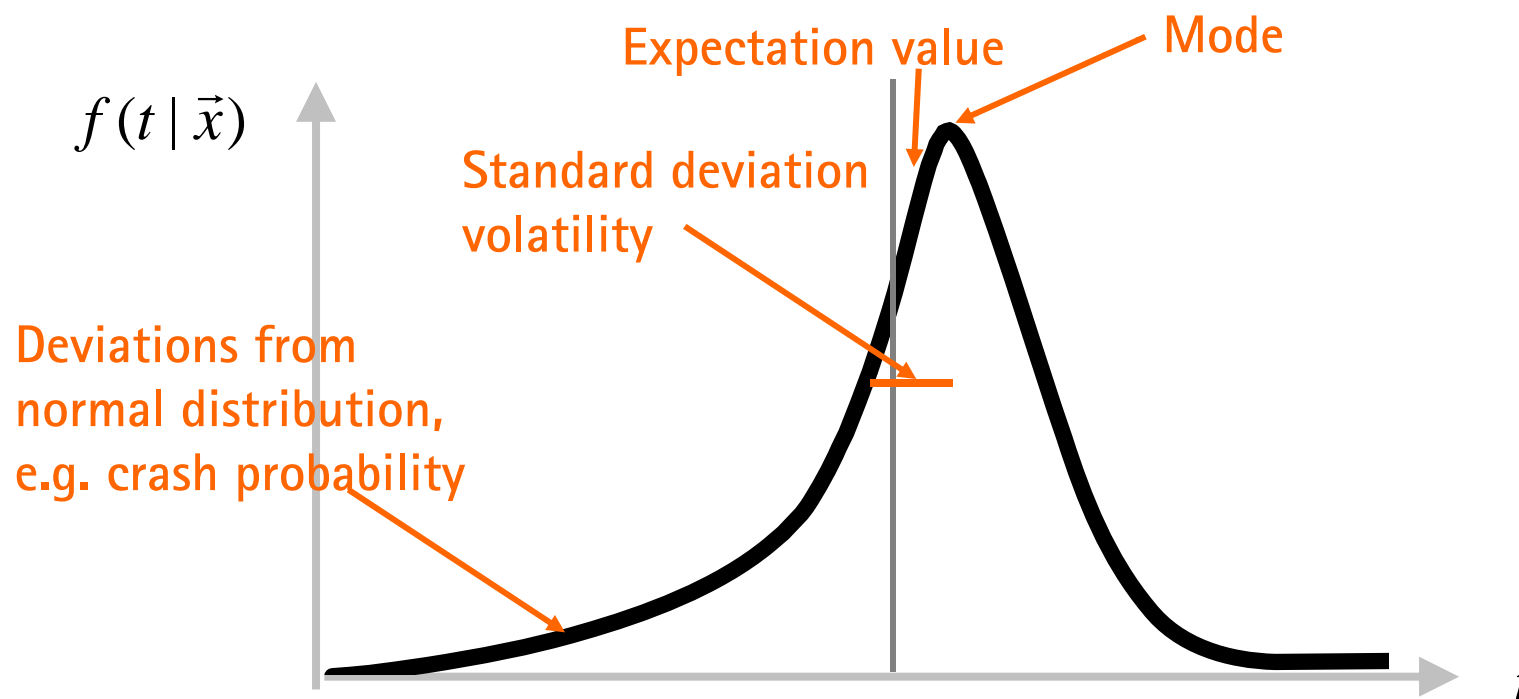
## Probability density for real valued targets:

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

## Examples:

- > Energy of an elementary particle  
(e.g a semileptonically decaying B meson with missing neutrino)
- > Q value of a decay
- > Lifetime of a decay
- > Price change of an equity or option
- > Company turnaround or earnings

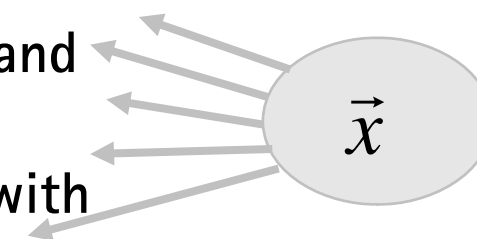
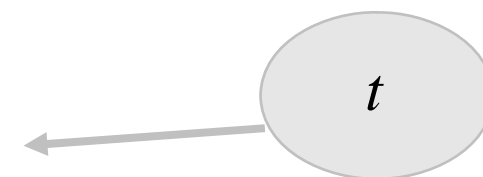
# Prediction of the complete probability distribution – event by event unfolding –



# Conditional probability densities in particle physics

What is the probability density  
of the true B momentum  
in this semileptonic B candidate event  
taken with the CDF II detector

with these n tracks with those momenta and  
rapidities in the hemisphere,  
which are forming this secondary vertex with  
this decay length and probability, this invariant  
mass and transverse momentum, this lepton  
information, this missing transverse momentum,  
this difference in Phi and Theta between  
momentum sum and vertex topology, etc pp

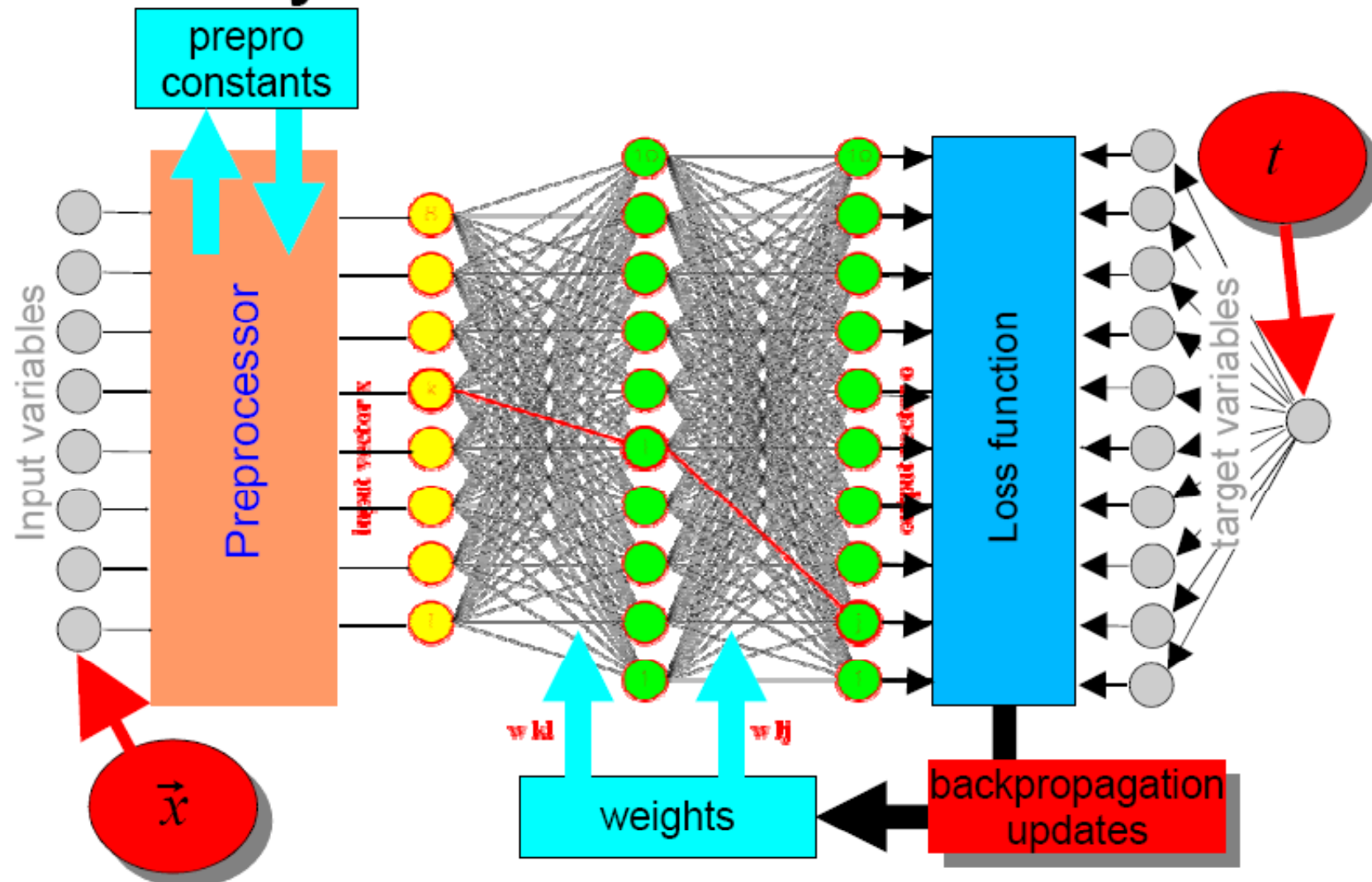


$$f(t | \vec{x})$$

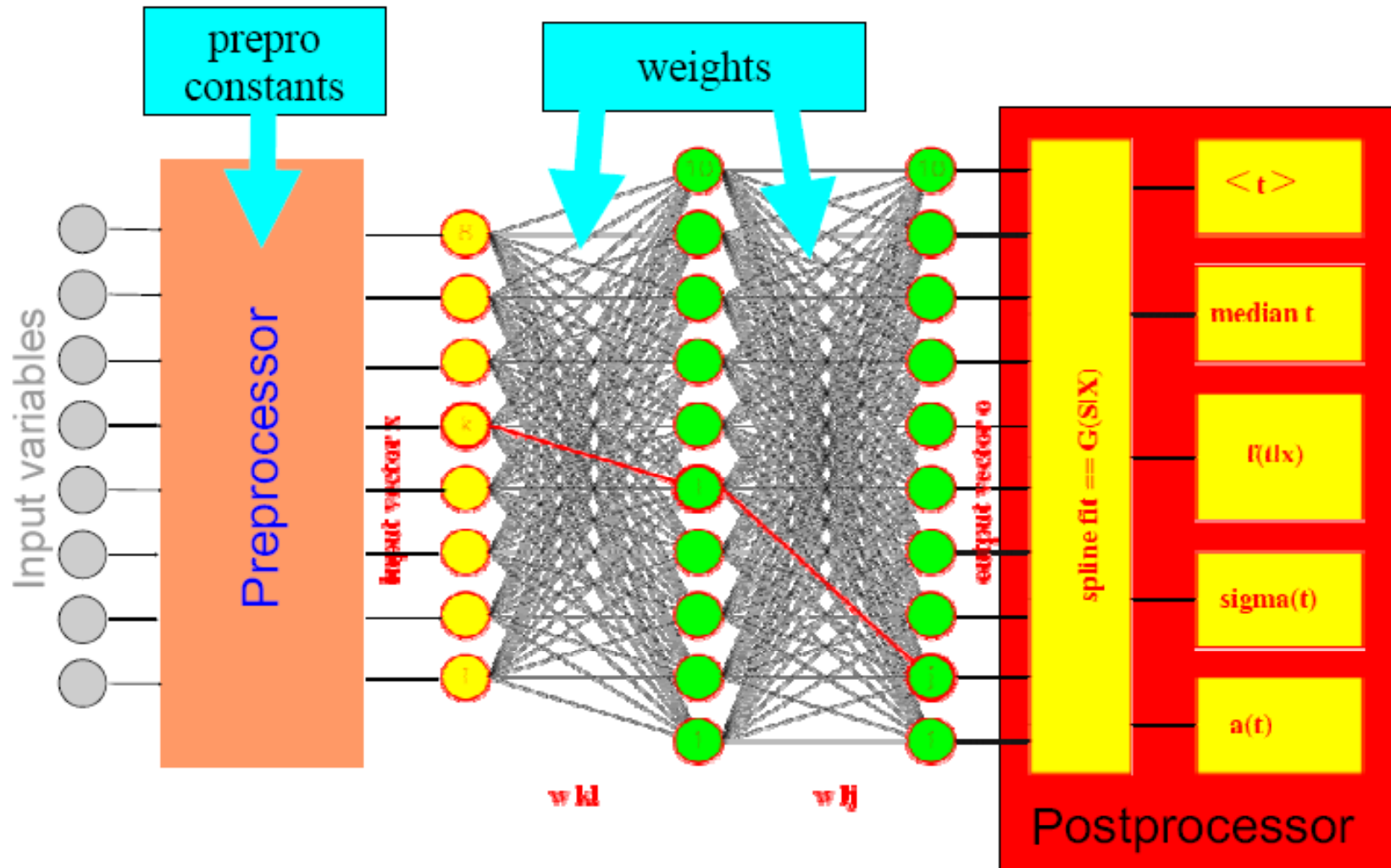
## <phi-t> NeuroBayes<sup>®</sup>

- > is based on neural 2nd generation algorithms, Bayesian regularisation, optimised preprocessing with transformations and decorrelation of input variables and linear correlation to output.
- > learns extremely fast due to 2nd order methods and 0-iteration mode
- > can train with weights and background subtraction
- > is extremely robust against outliers
- > 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
- > has advanced boost and cross validation features
- > is steadily further developed

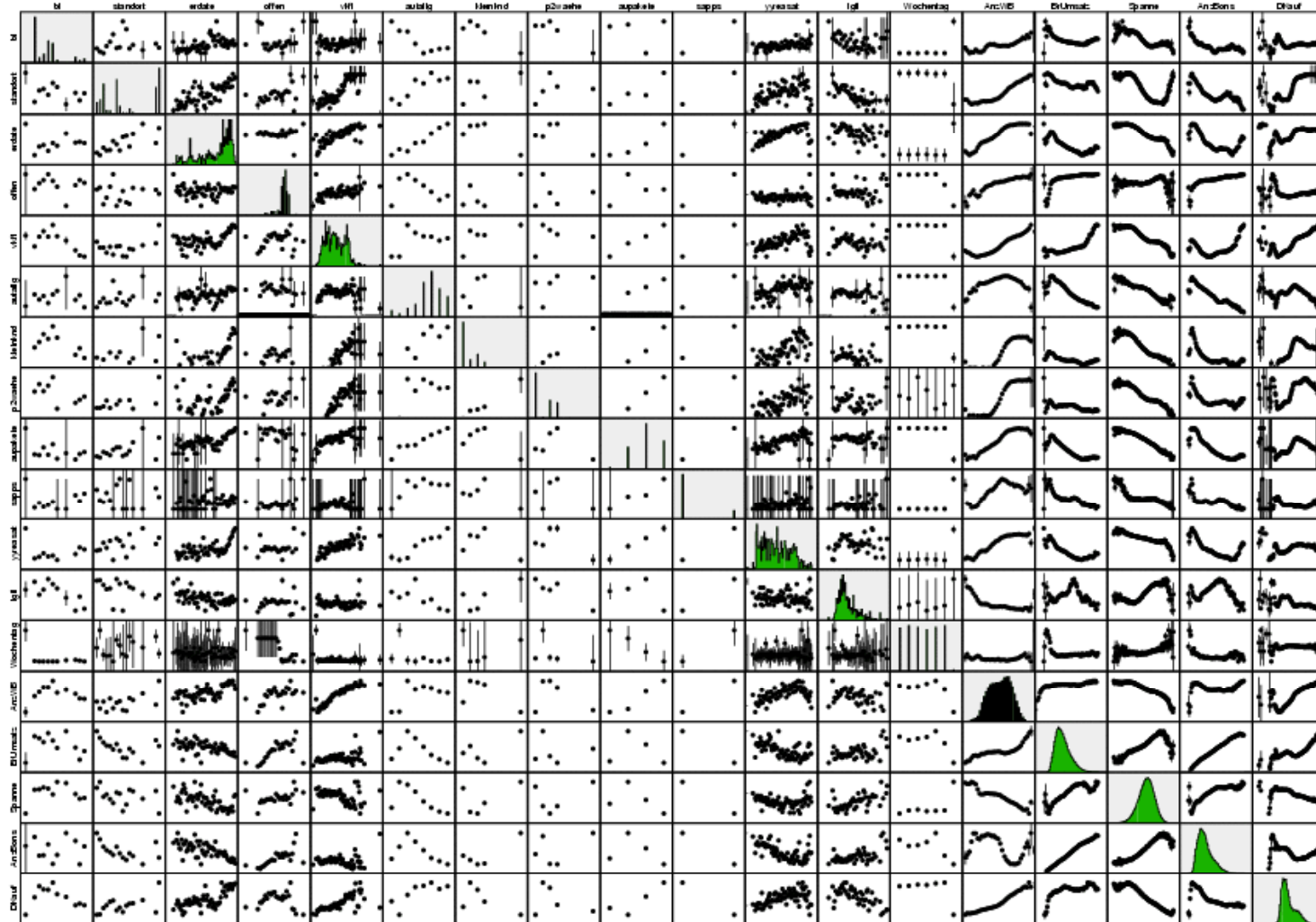
# NeuroBayes Network architecture: Teacher



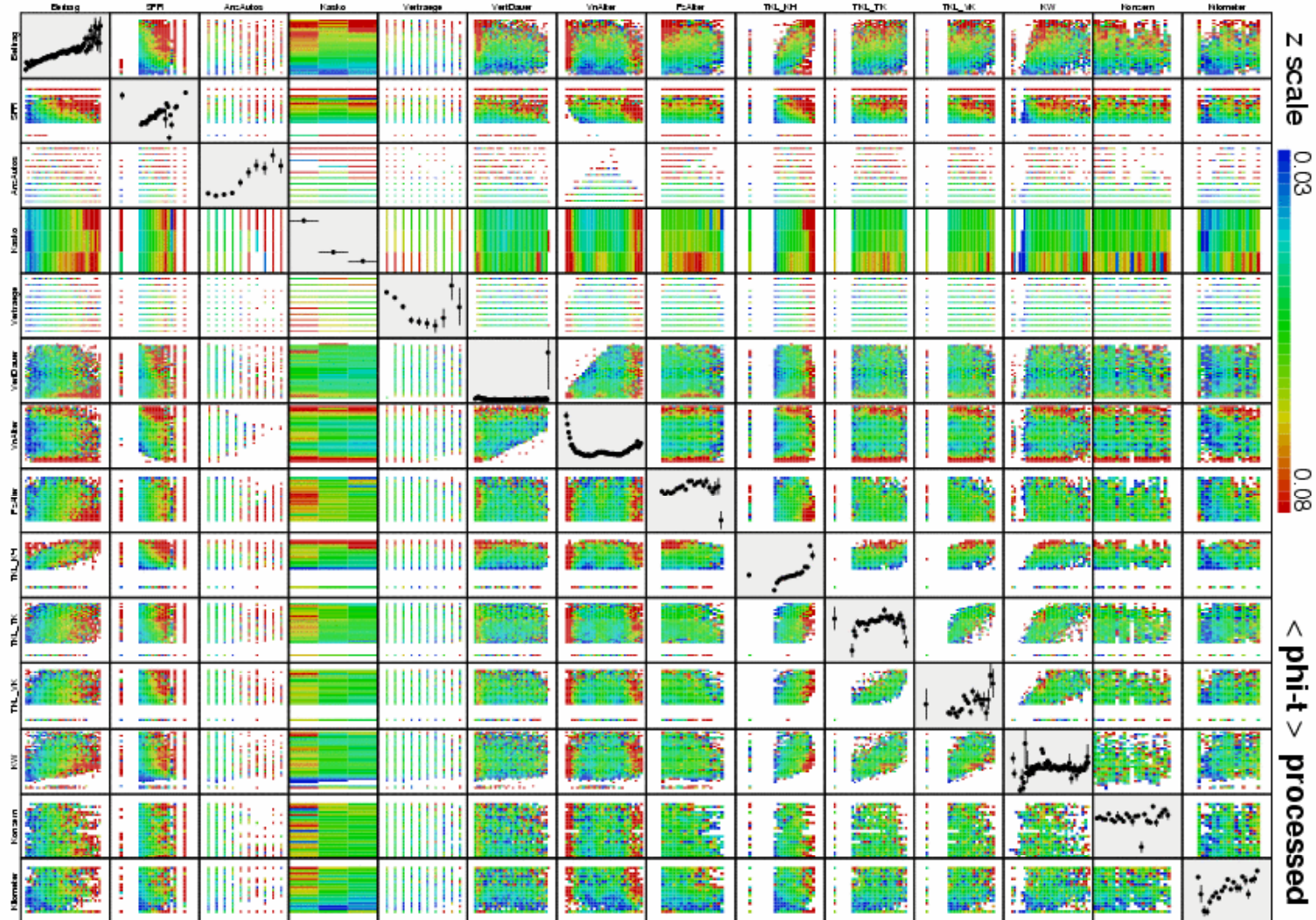
# NeuroBayes network architecture: Expert



# Ramler-plot (extended correlation matrix)

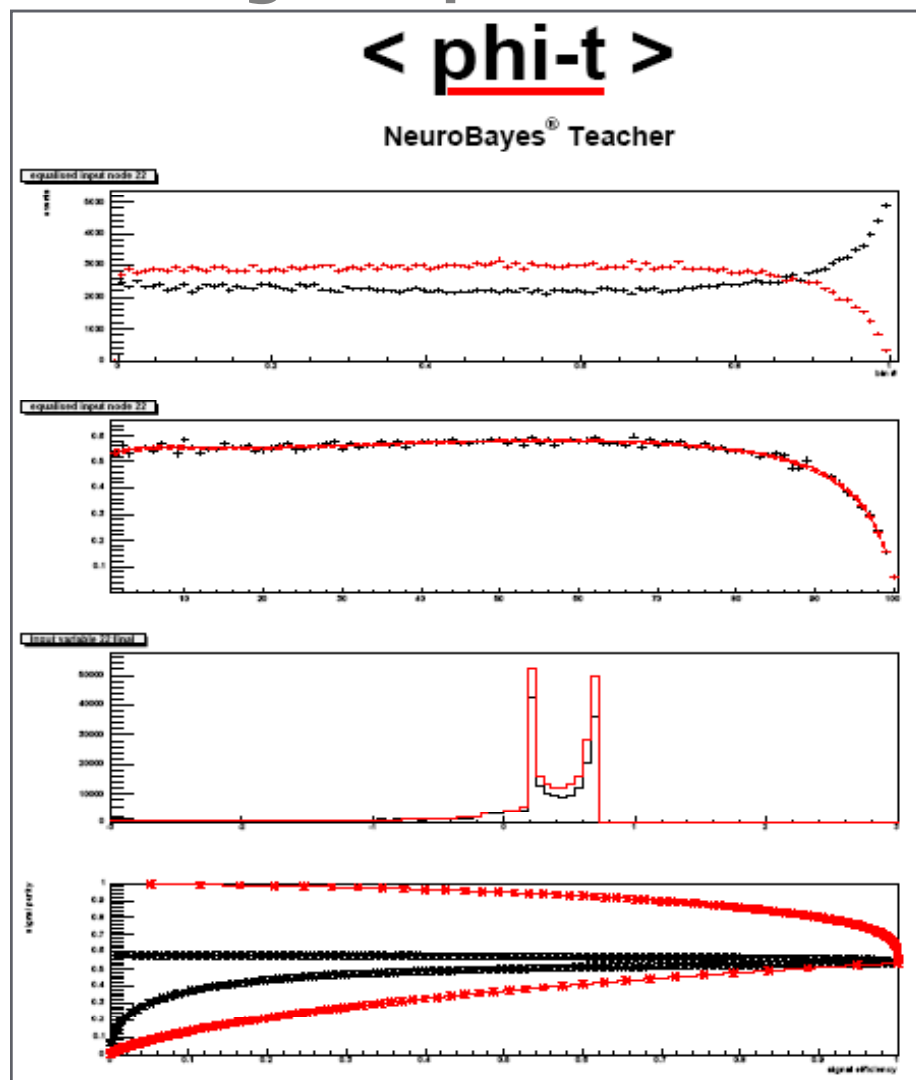


# Ramler-II-plot (visualize correlation to target)



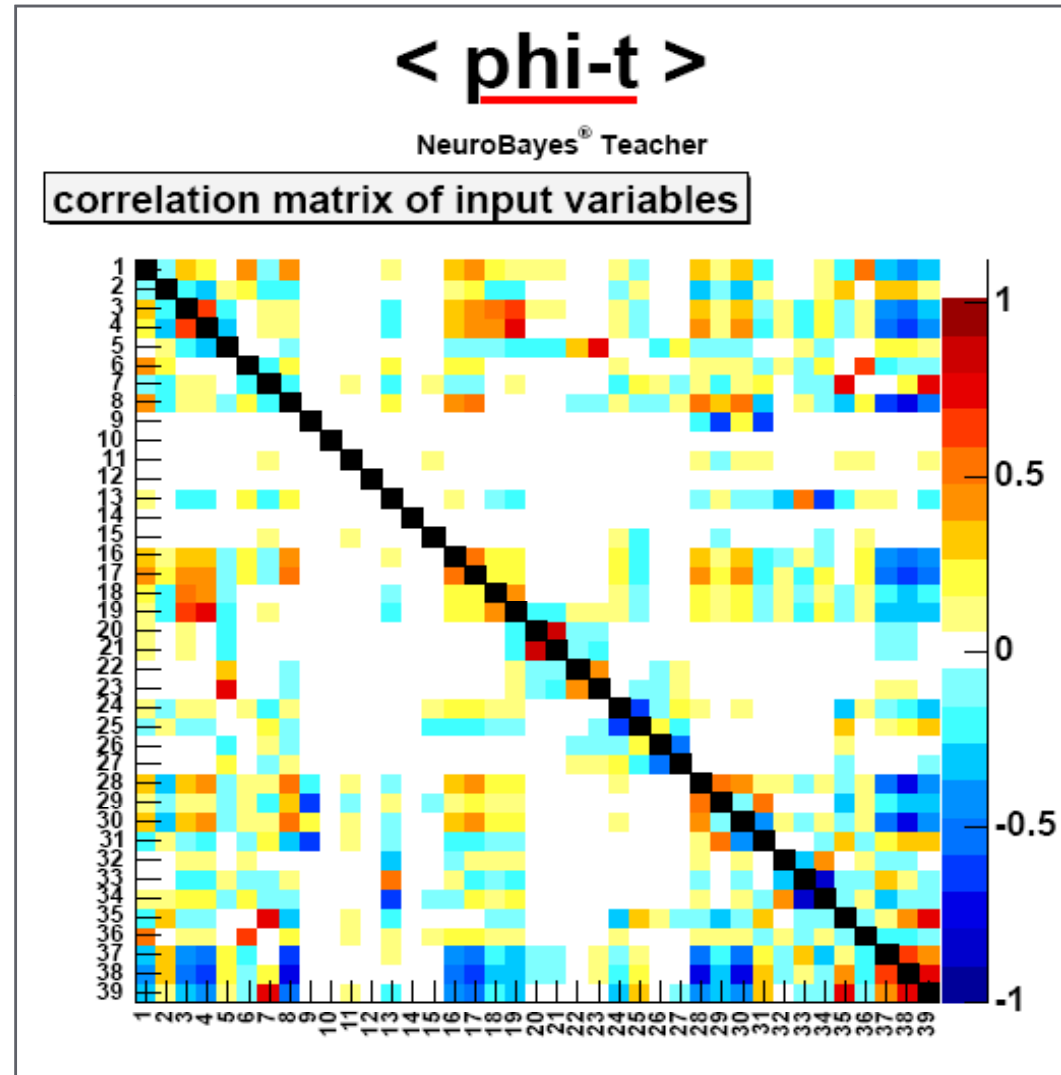


# Visualisation of single input-variables

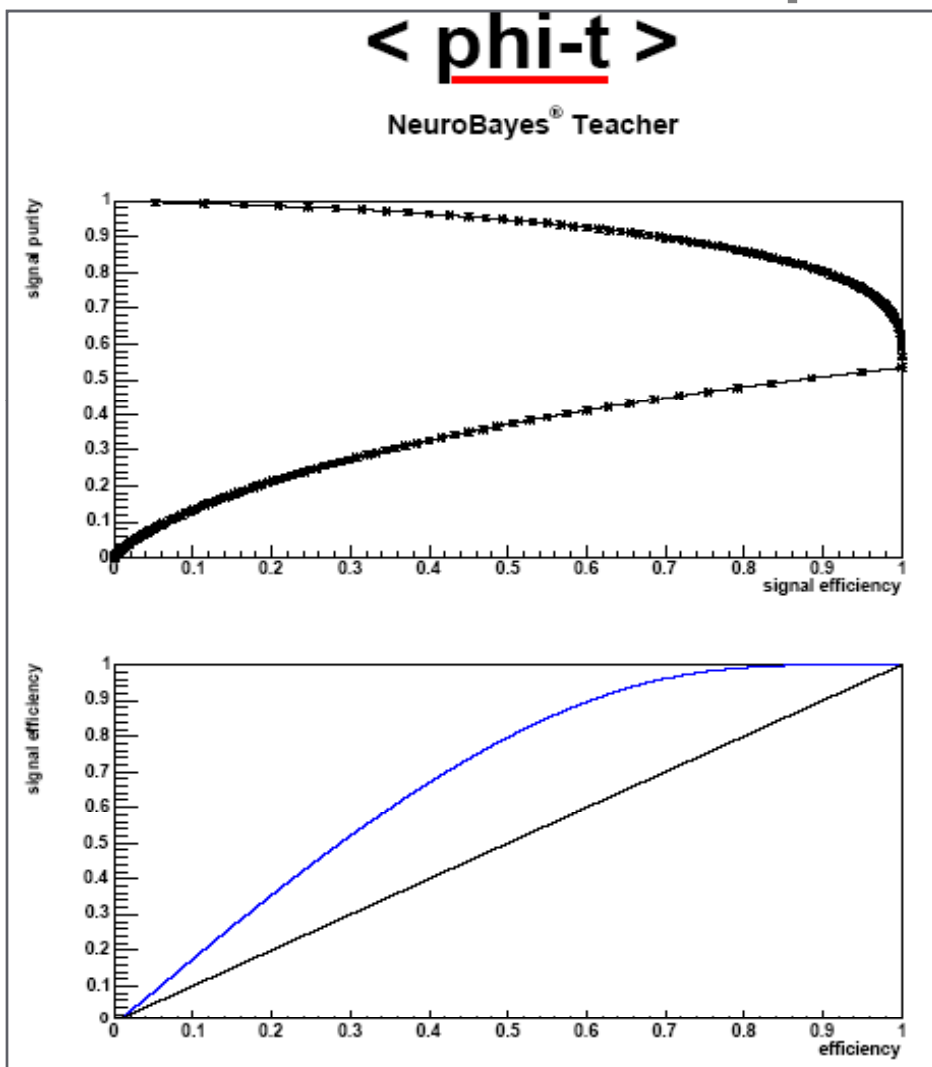


# Visualisation of correlation matrix

Variable 1: Training target



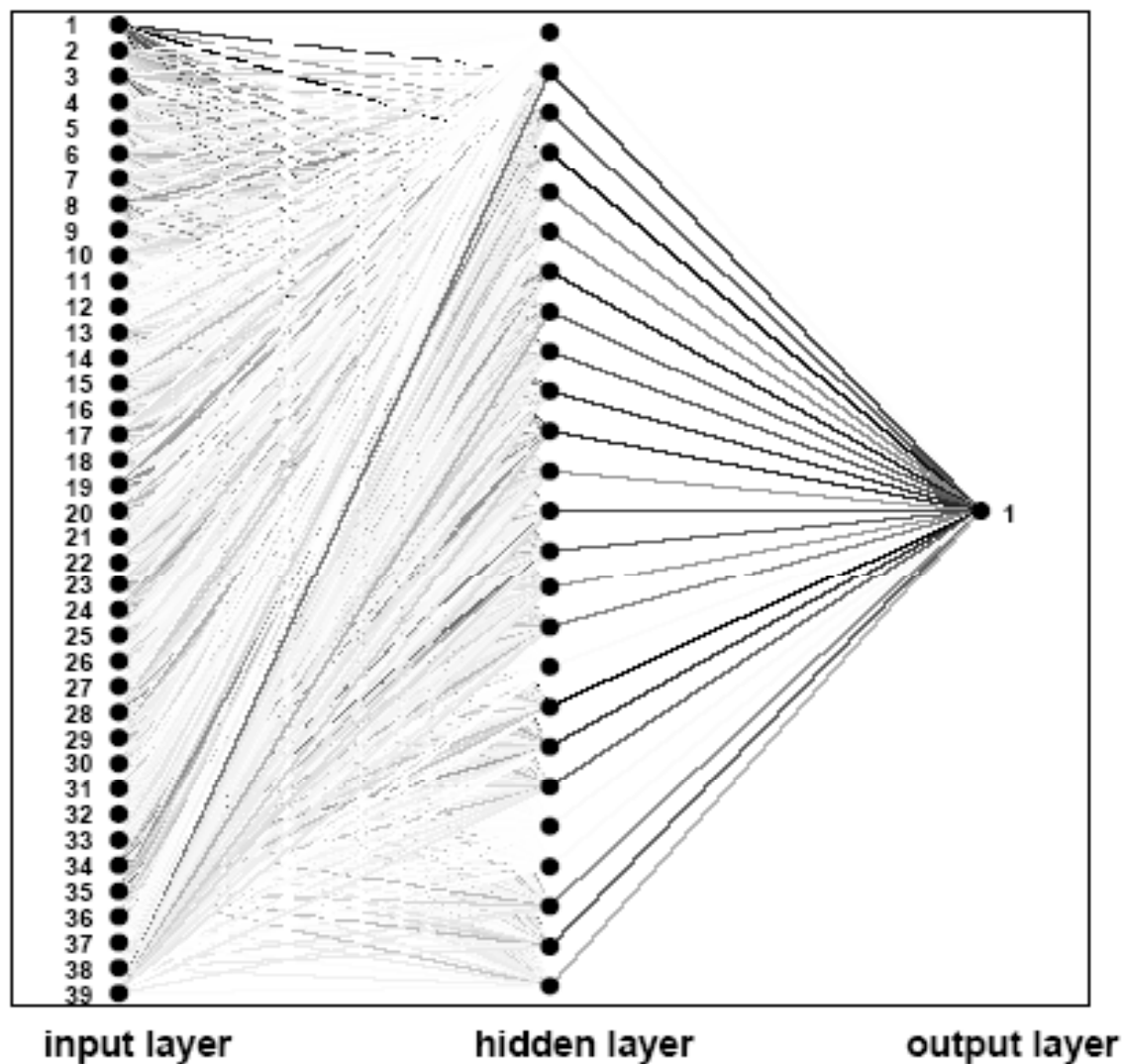
# Visualisation of network performance



Purity vs. efficiency

Signal-efficiency vs.  
total efficiency  
(Lift chart)

# Visualisation of NeuroBayes network topology



## Successful applications in High Energy Physics:

### at CDF II:

Electron ID, muon ID, kaon/proton ID

Optimisation of resonance reconstruction in many channels

(X, Y, D, D<sub>s</sub>, D<sub>s</sub><sup>\*\*</sup>, B, B<sub>s</sub>, B<sup>\*\*</sup>, B<sub>s</sub><sup>\*\*</sup>)

Spin parity analysis of X(3182)

Inclusion of NB output in likelihood fits

B-tagging for high pt physics (top, Higgs, etc.)

Single top quark production search

Higgs search

B-Flavour Tagging for mixing analyses (new combined tagging)

B<sub>0</sub>, B<sub>s</sub>-lifetime, DG, mixing, CP violation

## Further NeuroBayes<sup>®</sup> applications in HEP:

### AMS:

Particle ID design studies

### CMS (ongoing) :

B-tagging

### Belle (ongoing) :

Continuum suppression

### H1 (ongoing) :

Hadron calorimeter energy resolution optimisation

More than 35 diploma and Ph.D. theses...

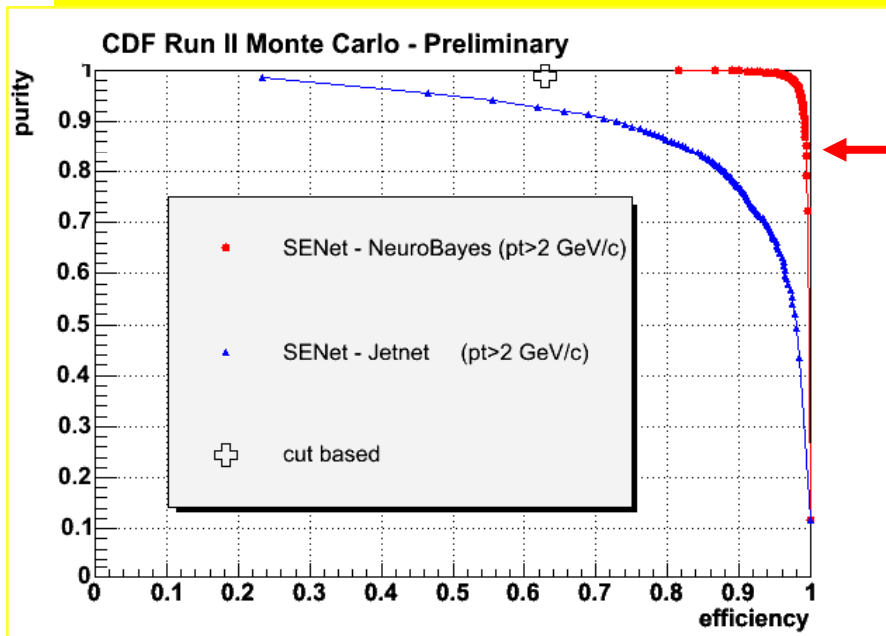
from experiments DELPHI, CDF II, AMS and CMS  
used NeuroBayes<sup>®</sup> or predecessors very successfully.

Many of these can be found at  
[www.phi-t.de](http://www.phi-t.de) → Wissenschaft → NeuroBayes

Talks about NeuroBayes<sup>®</sup> and applications:  
[www-ekp.physik.uni-karlsruhe.de/~feindt](http://www-ekp.physik.uni-karlsruhe.de/~feindt) → Forschung

# Just a few examples...

NeuroBayes soft electron identification for CDF II  
(Ulrich Kerzel, Michael Milnik, M.F.)



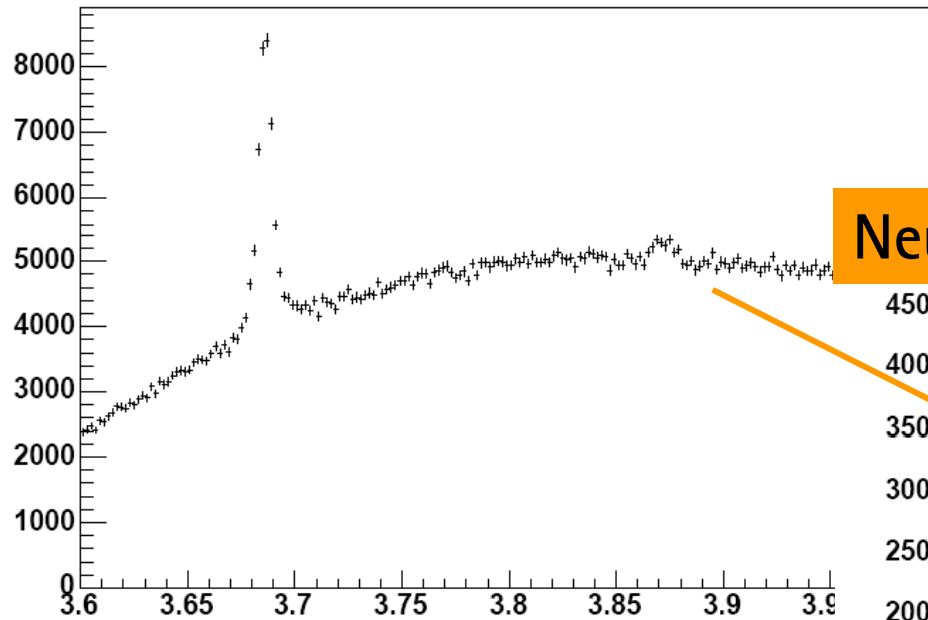
Thesis U. Kerzel:  
on basis of Soft Electron Collection  
(much more efficient than  
cut selection  
or **JetNet with same inputs**)  
- after clever preprocessing by hand  
and careful learning parameter  
choice this could also be as good  
as NeuroBayes®



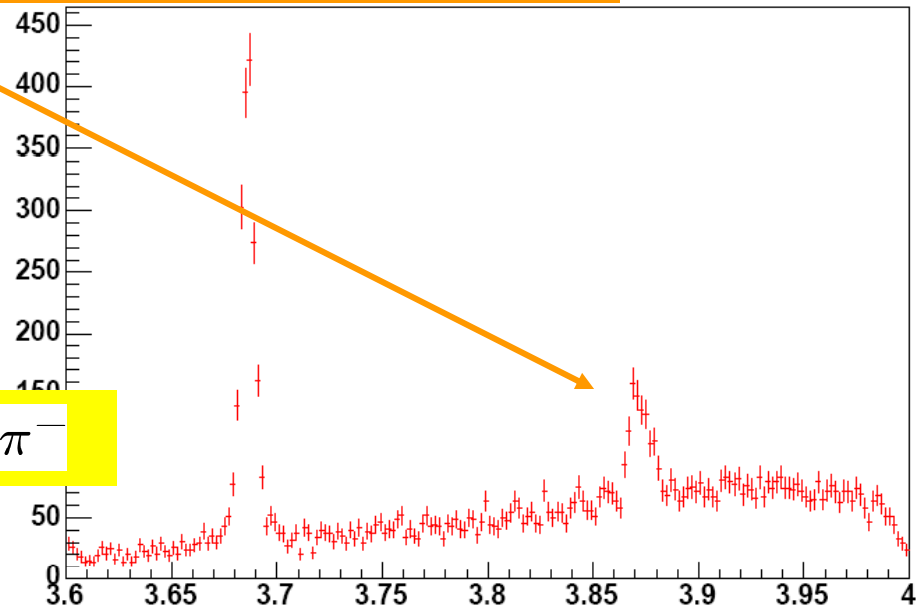
# Just a few examples...



mass (J/Psi Pi Pi)



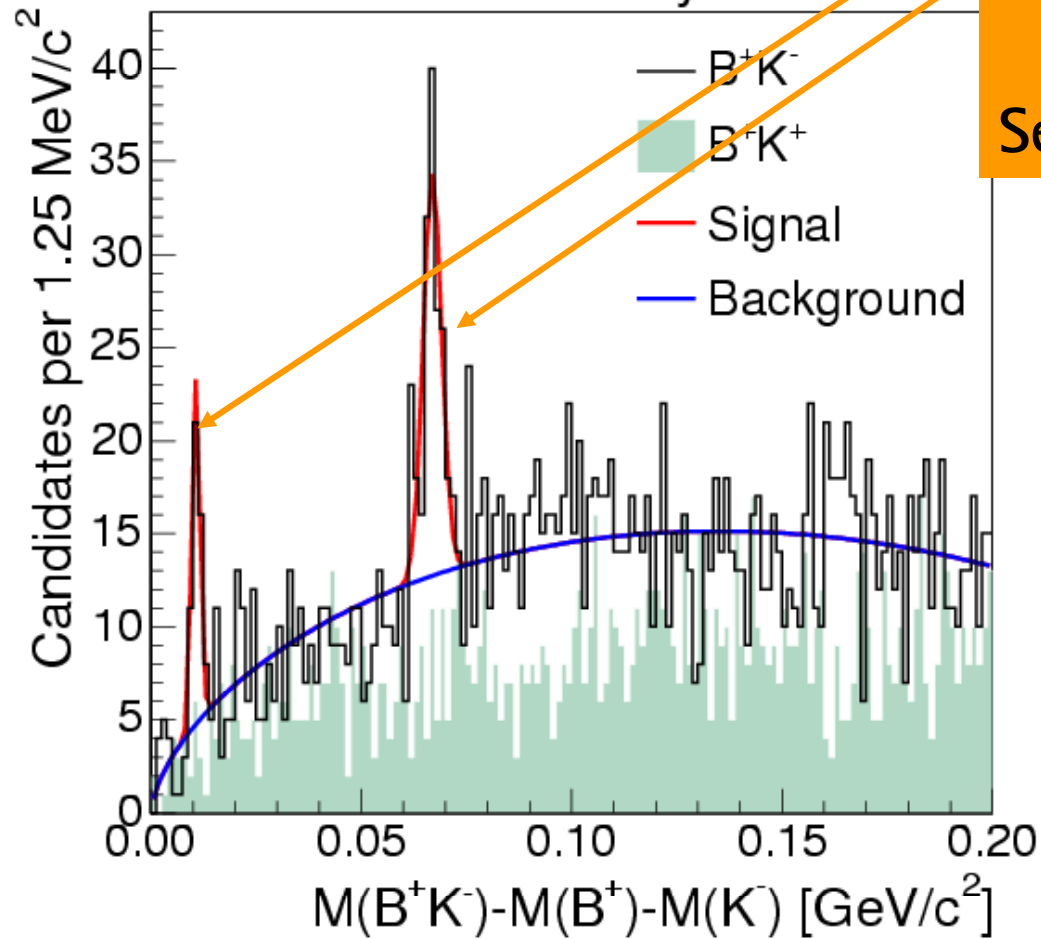
NeuroBayes<sup>®</sup> selection



$\psi(2S)$  and  $X(3872)$  in  $J/\psi\pi^+\pi^-$

# Just a few examples...

CDF Run 2 Preliminary 1.0 fb<sup>-1</sup>



First observation of B<sub>s1</sub> and most precise of B<sub>s2</sub>\*

Selection using NeuroBayes®





# Successful in competition with other data-mining- methods

World's largest students  
competition:

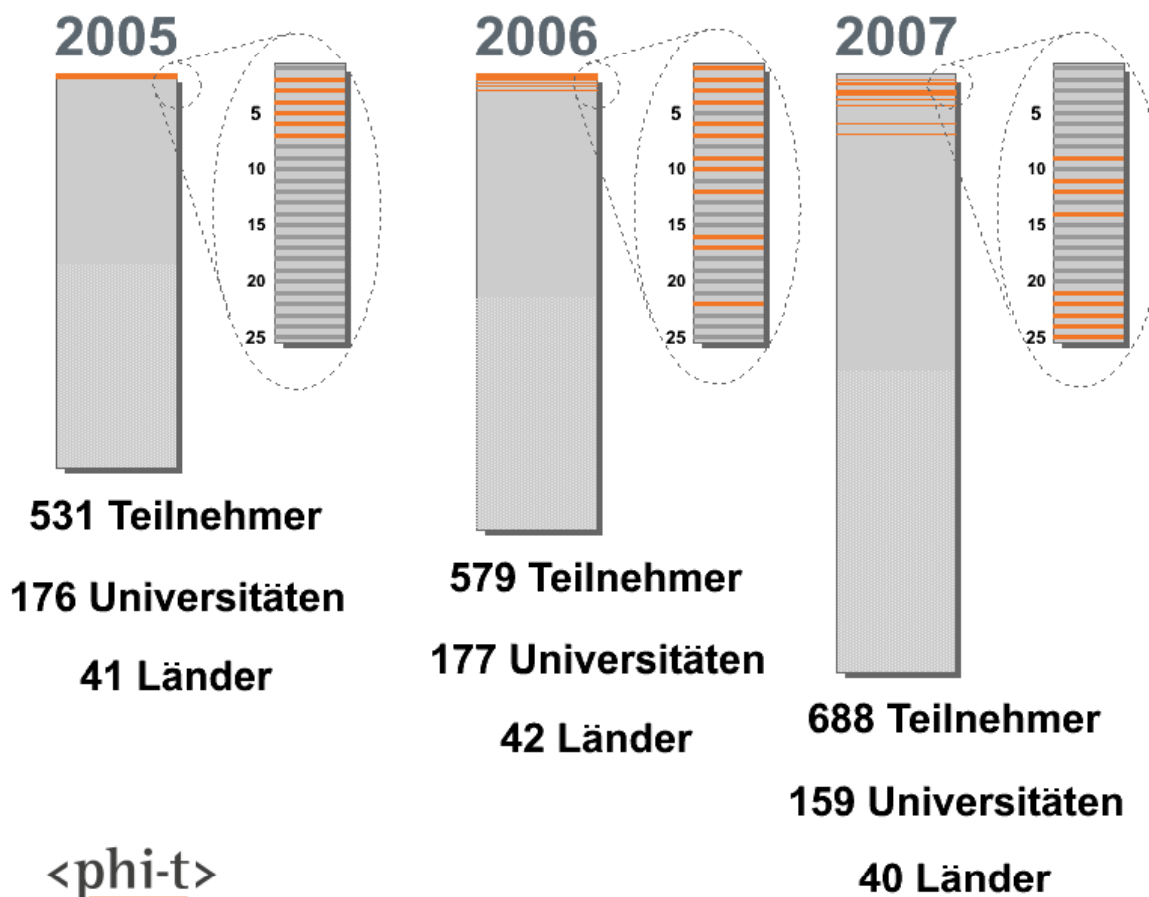
**Data-Mining-Cup**

**2005:** Fraud detection  
in internet trading

**2006:** price prediction  
in ebay auctions

**2007:** coupon redemption  
prediction

## Platzierungen von NeuroBayes<sup>®</sup>-Lösungen beim Data-Mining-Cup



# Phi-T: Applications of NeuroBayes<sup>®</sup> in Economy

## > Medicine and Pharma research

e.g. effects and undesirable effects of drugs  
early tumor recognition

## > Banks

e.g. credit-scoring (Basel II), finance time series  
prediction, valuation of derivatives, risk minimised  
trading strategies, client valuation

## > Insurances

e.g. risk and cost prediction for individual clients,  
probability of contract cancellation, fraud recognition,  
justice in tariffs

## > Trading chain stores:

turnover prognosis for individual articles/stores

Necessary prerequisite:

Historic or simulated data must be available!

Physics Information Technologies

# Prophetin des Knalls

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

Von Rainer Kayser

Michael Feindt ist Physiker. Ein Physiker auf Abwegen: Denn der Teilchenexperte der Universität Karlsruhe hilft deutschen Versicherungen, Tarifstrukturen zu optimieren. Schließlich haben Teilchenbeschleuniger und Versicherungen etwas gemeinsam: Bei beiden fällt eine fast unüberschaubare Menge von Daten an. In diesen Datenbergen schlummern unter Umständen Schätze: Spuren von Elementarteilchen – oder Hinweise auf besonders risikofreudige Autofahrer.

Warum also, fragte sich Feindt, nicht in beiden Fällen dieselben mathematischen Verfahren zur Schatzsuche verwenden? Seit über zwei Jahren analysiert der Physiker nun mit seiner Firma Phi-T erfolgreich die Daten Fremder: Er betreut mehrere Kunden, darunter die Badischen Versicherungen, die Drogeriekette DM und die Bausparkasse Schwäbisch Hall. Die Badischen Versicherungen sind die Ersten, die die Ergebnisse des Feindt'schen Analyseverfahrens umsetzen: Die Karlsruher Versicherungsgruppe veränderte ihre Tarifstruktur für Autofahrer.

Das Herzstück des Verfahrens ist das Programm NeuroBayes, ein neuronales Netz. Darunter versteht man eine Software, die nach dem Vorbild der vernetzten Nervenzellen im Gehirn arbeitet. Sie ist lernfähig und passt ihre Verarbeitungsprozesse ständig ihren Erfahrungen an.

Ein solches System kann komplexe Aufgaben effektiver bearbeiten als andere Computerprogramme. Bislang wurden neuronale Systeme in der Bildverarbeitung, in Expertensystemen und bei der Prozessoptimierung eingesetzt.

NeuroBayes sucht in Daten nach Strukturen und Zusammenhängen. Dabei ist es egal, ob es sich um Messdaten der Teilchenphysik handelt oder um Datenmaterial von Versicherungen. „In das Verfahren sind 20 Jahre Analyse-Erfahrung bei Teilchen-Experimenten und das gesammelte Wissen von 2000 Physikern aus aller Welt eingeflossen“, sagt Feindt.

**„Wir haben bereits neue Tarife eingeführt“**

Heinz Ohnmacht, Badische Versicherungen

Das Ziel des Verfahrens ist, bislang unbekannte Muster in Daten aufzuspüren: Muster, die auf ein neues subatomares Teilchen hindeuten oder auf bisher unerkannte Zusammenhänge zwischen Autofahrern und Unfallrisiko. Doch neuronale Netze können bei der Analyse genauso irren wie unser Gehirn, das – etwa bei optischen Täuschungen – dazu neigt, Strukturen auch dort zu erkennen, wo keine sind. Mit speziellen statistischen Verfahren verhindert Feindt deshalb, dass NeuroBayes überreagiert und Zusammenhänge sieht wo nur der Zufall regiert.

Schon länger differenzieren Kfz-Versicherer ihre Tarife nach Fahrzeugtypen und Regionalklassen. NeuroBayes optimiert diese Zusammenhänge

zwischen Auto, Fahrer, Wohnort und Fahrverhalten auf. Aufgrund der Analyse von Phi-T haben wird bereits eine feinere Tarifstruktur bei jungen Fahrern eingeführt“, sagt Heinz Ohnmacht, der bei den Badischen Versicherungen für den Kfz-sich beispielsweise geizigt, dass das Unfallrisiko weiblicher Fahrer erfahrungsbildig geringer ist als das männlicher Verkehrsneulinge.

Eine zuvor unterschätzte Rolle spielt auch das Alter des Autos: „Die jungen Fahrer passen auf einen neuen BMW einfach besser auf als auf einen alten Audi“, sagt der Versicherungsexperte. Zudem spielt auch eine Rolle, wie alt das Auto zum Zeitpunkt des Kaufs ist: „Ein Auto, das mit dem Fahrer gemeinsam altert, wird mehr geschätzt als ein bereits alt gekauftes“, sagt Ohnmacht. Der Versicherungsexperte hofft, dass man bald für jeden Fahrer ein ganz individuelles Unfallrisiko berechnen kann – und anschließend entsprechend individualisierte Prämien einführen kann. Wobei das die Frage aufwirft, welchen Sinn eine Versicherung noch hat, wenn jeder nur sein eigenes Risiko versichert: Könnte da nicht jeder seinen Schaden selbst bezahlen und so die Kosten der Versicherung sparen?

## Clevere Netze

Im Gehirn sind die Nervenzellen (Neuronen) miteinander vernetzt. Dadurch können sie Informationen austauschen. Trifft ein bestimmter Input häufiger ein, verändert sich die Zelle, sie lernt.

Neuronale Netze sind Computerprogramme mit speziellen mathematischen Algorithmen. Sie sind nach dem Vorbild der Nervenzellen im Gehirn aufgebaut. Sie sind lernfähig und passen ihre Verarbeitungsprozesse ständig ihren Erfahrungen an.

NeuroBayes ist ein neuronales Netz, ursprünglich entworfen, um neue Elementarteilchen zu entdecken. Doch es kann auch kleinste statistische Zusammenhänge erkennen. Diese Fähigkeit kann Versicherungen helfen, Risiken einzuschätzen.

„Nicht ganz unberechtigt“ sei dieser Einwand, gesteht Ohnmacht ein, „aber das gilt nur für kleine Schäden. Große Schäden können natürlich auch weiterhin nur durch eine große Versicherungsgemeinschaft aufgefangen werden.“ Zudem liefere das Verfahren nur Wahrscheinlichkeiten: „Wen es tatsächlich trifft, ist letztlich Zufall.“

Ein Zufall war es auch, der Physiker Feindt auf kommerzielle Wege führte. „Ich habe mit meinen Aktien Verluste erlitten“, so der Forscher, „deshalb kam ich auf die Idee, mit meiner Methode Aktienkurse vorherzusagen.“ Ob NeuroBayes den Börsenkurs tatsächlich prophezeien kann, ist im Internet überprüfbar: www.phi-t.de/mousegame.

Das Programm braucht Alter des Fahrers und des Wagens – dann ist der Autocrash fast vorhersagbar



Individual risk prognoses  
for car insurances:

Accident probability  
Cost probability distribution  
Large damage prognosis  
Contract cancellation prob.

→ very successful at

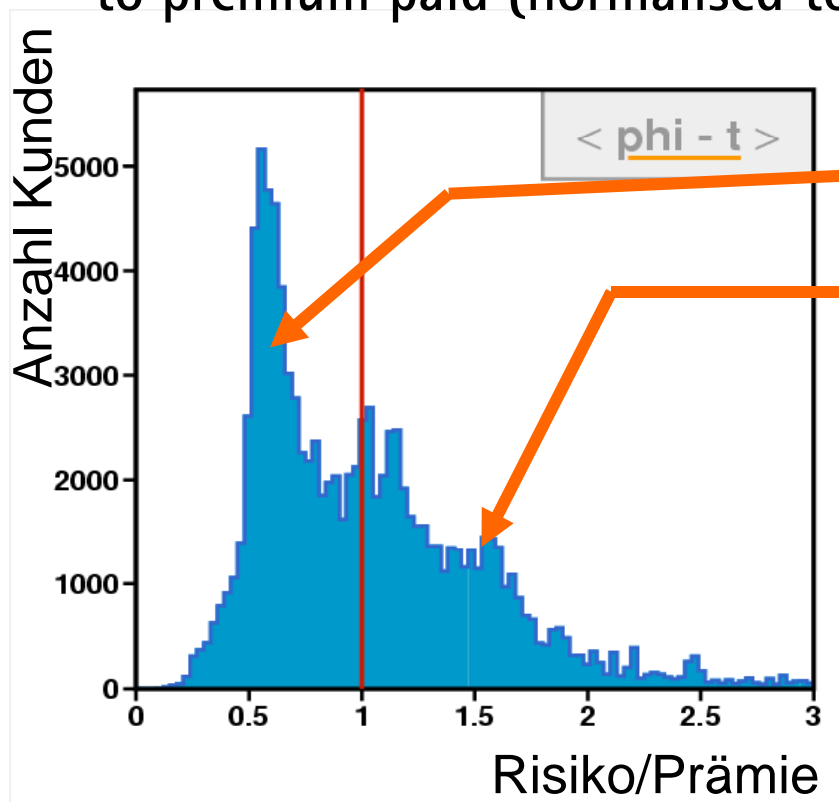
Badisch gut versichert.

BGVA

Physics

## The "unjustice" of insurance premiums

Ratio of the accident risk calculated using NeuroBayes<sup>®</sup> to premium paid (normalised to same total premium sum):



The majority of customers (with low risk) are paying too much.

Less than half of the customers (with larger risk) do not pay enough, some by far not enough.

These are currently subsidised by the more careful customers.

Physics Information Technologies



„Wenn man alles auf physikalische Gesetzmäßigkeiten zurückführen würde, wäre das eine Abbildung mit inadäquaten Mitteln, so als ob man eine Beethoven-Symphonie als Luftdruckkurve darstellte.“

Albert Einstein

# Physiker helfen Versicherern

## Statistische Verfahren aus der Elementarteilchen-Forschung berechnen aus komplexen Daten maßgeschneiderte Tarife

VON ANTONIA ROTGER

**Karlsruhe** – Es gibt Dinge, die sind perfekt berechenbar, und Dinge, die vollkommen zufällig sind. Doch zwischen der Genauigkeit, mit der wir den Sonnenaufgang vorhersagen und dem reinen Zufall, der den Reiz des Lottospiels ausmacht, erstreckt sich ein riesiges Zwischenreich: Wetter, Gesundheit oder Wirtschaftsentwicklung sind weder vom schieren Zufall gesteuert, noch im Detail für lange Zeit vorhersagbar. Und Wissen aus Unsicherheit zu ernten, ist eine hohe Kunst, die ihre Wurzeln in der Statistik und Wahrscheinlichkeitsrechnung hat. Einer der Experten heißt Michael Feindt, Professor an der Universität Karlsruhe. Als Teilchenphysiker sucht er am CERN und am Fermilab in Chicago im scheinbar sinnlosen Rauschen nach Teilchen und Antiteilchen, die aus dem Nichts für einen Augenblick entstehen.

Aus dem vorhandenen „Pool“ an mathematischen Werkzeugen hat

er aber auch ein Rechenverfahren entwickelt, das in dem Datenberg von Banken, Versicherungen oder Gesundheitsökonomien spannende Zusammenhänge finden kann. Sein Programm mit dem Namen NeuroBayes funktioniert dabei offenbar noch etwas besser als viele andere Datamining-Verfahren. Beim letzten Internationalen DataMining Cup 2005, bei dem 531 Teilnehmer aus 41 Ländern ihre Rechenverfahren gegeneinander ausspielten, haben seine zwei Studentinnen und vier Studenten die ersten Plätze abgeräumt.

Im Nebenberuf ist der Teilchenphysiker und Hochschullehrer nun auch Firmengründer, sein kleines Unternehmen Physics Information Technologies oder Phi-T arbeitet für die Versicherungswirtschaft, prognostiziert den Einkaufsbedarf für eine Drogeriekette oder berät Banken, die das Kreditrisiko von bereits verschuldeten Kunden einschätzen möchten.

Das Programm NeuroBayes besteht aus drei Schritten: Zunächst

werden die Daten mit einem Programm vorbereitet, das sie nach ihrer Bedeutung gewichtet. Hier kommt die Erfahrung ins Spiel: Welche Parameter haben sich schon früher als entscheidend erwiesen, welche Informationen eher

die Information aus der Unfallstatistik einer Pkw-Versicherung, daß der einzige versicherte Lamborghini-Fahrer im vergangenen Jahr einen Unfall gebaut habe. Daraus zu schließen, daß nun auch im nächsten Jahr 100 Prozent der

„Wir wollen auch kleine Zusammenhänge sehen, die sonst durch das Raster fallen.“

Michael Feindt, Universität Karlsruhe

nebensächlich? Dann errechnet das Hauptprogramm, ein sogenanntes Neuronales Netz, anhand der Daten aus Vergangenheit und Gegenwart, welche Ereignisse ursächlich miteinander zusammenhängen und welche Fakten nur rein zufällig gemeinsam auftreten. Dabei sind Korrekturen sowie das systematische Vergessen von „unwichtigen“ Zusammenhängen eingebaut. Unwichtig, weil statistisch nicht aussagefähig, wäre zum Beispiel

Lamborghini Unfälle verursachen, wäre vollkommener Unsinn. Das NeuroBayes-Programm sortiert solche Informationen zuverlässig aus. Erst nach diesen Schritten geht es an die Prognose. „Es ist ganz wichtig, klar zwischen Wissen und Nichtwissen zu unterscheiden“, meint Feindt. Und die Experteneinschätzung, die ganz zu Beginn einfließt, kann am Ende korrigiert werden, falls nötig. „Wir wollen auch kleine Zusammenhän-

ge sehen, die sonst durch das Raster fallen“, sagt der Physiker. Die gängigen Programme suchen vor allem nach „Schubladen“, in die die Kunden einsortiert werden können. Diese Schubladen braucht das neue Programm nicht, „im Prinzip kann jeder Kunde individuell beurteilt werden“, sagt Feindt. Und nicht nur die Schadenswahrscheinlichkeit, sondern auch die Verteilung der Schadenshöhe läßt sich voraussagen. Junge Frauen verursachen zum Beispiel häufig Unfälle, allerdings meist mit geringem Schaden. Bei jungen Männern sind die Unfallraten ähnlich hoch, die Schadenshöhe aber meist deutlich größer. Dies nutzt nun die Badische Versicherung: „Aufgrund der Analyse haben wir eine feinere Tarifstruktur eingeführt, wir sind dadurch jetzt in der Lage, ausgewählte junge Fahrer günstiger zu versichern“, erklärt Heinz Ohnmacht von den Badischen Versicherungen.

Mathematisch gesehen ist es nicht trivial, für solche Probleme

eine optimale Lösung zu finden. Das sei wie eine Hügellandschaft, in der man sich leicht am vermeintlich tiefsten Punkt niederläßt, auch wenn es nur ein „lokales Minimum“ sei und es anderswo tiefere Gruben gebe.

Bei einem realistischen Problem kommen aber leicht eine Million Fälle zusammen, jeder einzelne davon ist durch bis zu 100 Parameter gekennzeichnet. Das NeuroBayes rechnet daran nur sechs Stunden. Dabei werden bei jedem Durchlauf viele Verbindungen gekappt, das Neuronale Netz dünnt sich aus. Dieser Prozeß erinnert an den Reifungsprozeß im Gehirn während der Kleinkindzeit: Erst durch das Jäten von unwichtigen Verbindungen zwischen Nervenzellen kommen die Signale effizient ans Ziel. Feindt sieht noch viel Potential für bessere statistische Verfahren. Gerade in der Gesundheitsforschung wären genauere Analysen nützlich, zum Beispiel für die individuelle Kosten/Nutzen-Abwägung bei riskanten Therapien.

# Versandhandel: Physiker helfen Otto beim Sparen

**HAMBURG** – Was haben Elementarteilchenphysik und ein T-Shirt aus dem Versandhandel gemeinsam? Eine ganze Menge, sagt der Physikprofessor Michael Feindt. Der Spezialist von der Uni Karlsruhe hilft jetzt dem Hamburger Versandhandelsriesen Otto, seine Absatzprognosen zu verbessern. Der Konzern hat mit Feindts Firma Phi-T ein Gemeinschaftsunternehmen gegründet und verspricht sich da-

senden Produkten wie beispielsweise einem T-Shirt gefüttert. Farbe, Größe oder Marke fließen ebenso in die Analyse ein wie der Preis oder die Platzierung der Ware im Katalog. Anhand dieser Eingaben ist das System in der Lage, eine Vorhersage über die verkauften Stückzahlen der Ware zu erstellen. Das neuronale Netz lernt dabei selbstständig und kann auch bislang unbekannte Wechselwirkungen zwi-



ation Technologies

losing Meeting, CERN

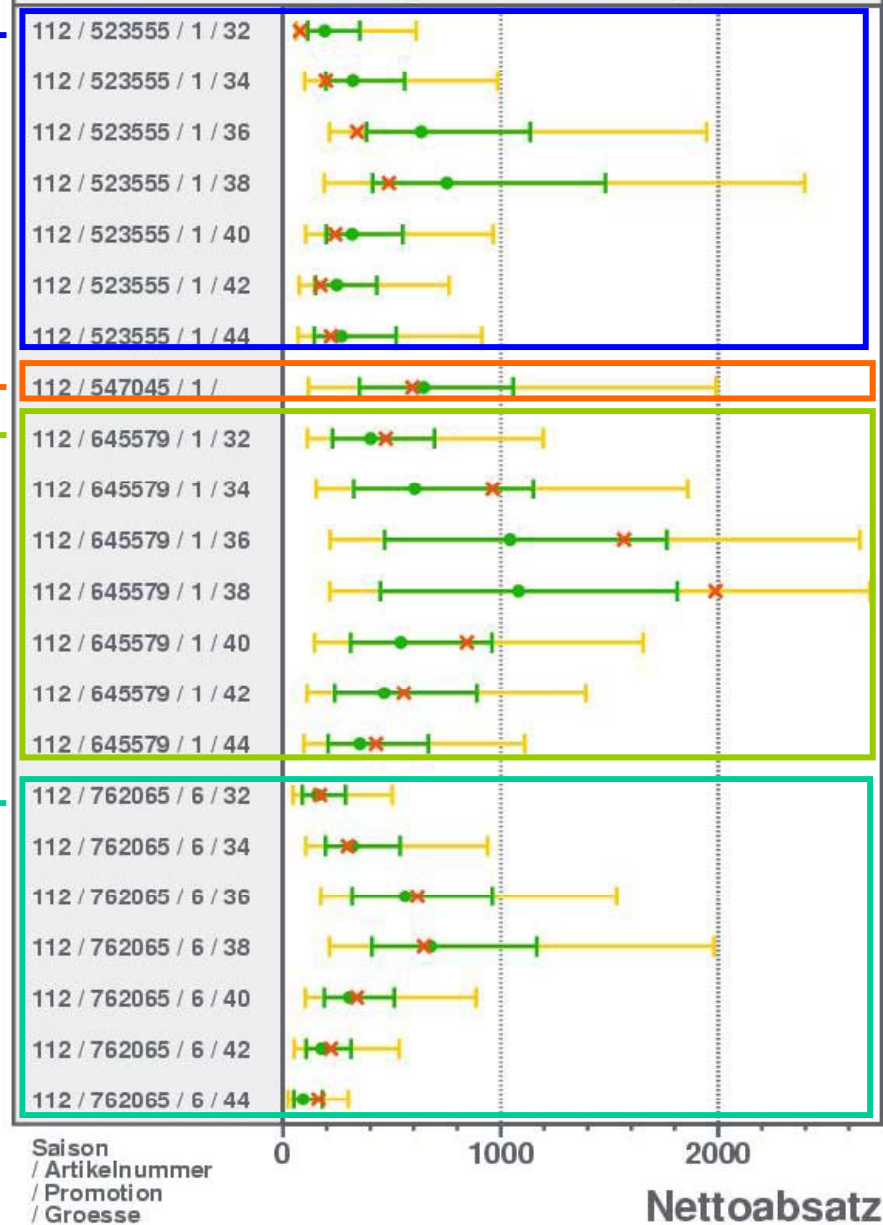
May 2008



# Turnover prognosis for mail order business



Denim-News: **Used de Luxe** lässt andere alt aussehen.



Otto... find' ich gut.  
Physics Infor

Ehe Quelle zum Beispiel im halbjährlich erscheinenden Hauptkatalog Modetrends abdrucken kann sind...

Eigenmarken, die schwer im Preis vergleichbar sind. Anders als überall erhältliche Marken wie Braun, Philips Go und LG. Hierfür verlangt der Versandhändler viel höhere Preise als die Konkurrenz bei Ebay und Amazon (siehe Preisvergleich Seite 31).

Katalogen teil, sondern eröffnen im Internet konkurrierende Shops. Nach einer Studie der Gesellschaft für Konsumforschung (GfK) ist der Wettbewerb im Einzelhandel europaweit nirgendwo so hart wie hierzulande. Der Bundesver-



Otto Versand in Hamburg versucht mit Rechenmodellen aus der Teilchenphysik, Bestellungen zu prognostizieren

### Physiker modernisieren Einkaufsmanagement bei Otto Bis ins kleinste Teilchen

Versandhandel und Teilchenphysik haben auf den ersten Blick nicht viel miteinander zu tun. Auf den zweiten aber eine ganze Menge. Das Versandgeschäft hat sich zu einer kom-

plexen Welt entwickelt. In der Teilchenphysik kann, werden Teilchen, x-fach kleiner als Atomkerne, aufeinandergeschossen. Bei jedem Experiment zeichnen die Computer riesige Datenmengen auf. Die Physiker müssen herausfinden, welche Daten zufällig sind und welche ein wiederkehrendes Muster

entdecken. In der Teilchenphysik sind zufällige Ausreißer von regelmäßigen Trends trennen. Der Vorteil aus Sicht der Physiker: „Zufälle gibt es wenige, fast alles folgt einem festen Muster“, sagt Feindt. In einem Testlauf haben die Karlsruher 2006 jeden Textilartikel des Otto-Katalogs und des Online-Angebots unter die Lupe genommen. Nach Farbvariante und Größe unterteilt, haben sie Schätzungen für die Bestellungen abgegeben. Fast immer lagen sie näher an den realen Daten als mit den Standard-Prognoseverfahren.

Die jahrelange Erfahrung eines Einkäufers wird die Software nicht ersetzen können. Aber Prognosen werden damit sicherer. Otto, seit Jahren intensiv um den Einsatz neuer Technologien bemüht, arbeitet eng mit Phi-T zusammen. Fallen die Tests weiter positiv aus, könnte Phi-T ab 2007 regelmäßig Schätzungen über die Gängigkeit von Sortimenten und Artikeln aus dem Badischen an die Elbe schicken. aj

plexen Welt entwickelt. In der Teilchenphysik kann, werden Teilchen, x-fach kleiner als Atomkerne, aufeinandergeschossen. Bei jedem Experiment zeichnen die Computer riesige Datenmengen auf. Die Physiker müssen herausfinden, welche Daten zufällig sind und welche ein wiederkehrendes Muster

Nichts anderes verlangt Otto, mit einem Gruppenumsatz von 14,6 Milliarden Euro im abgelaufenen Geschäftsjahr die Nummer 2 der Versandhandelsbranche nach Amazon, von Phi-T. Das Karlsruher IT-Unternehmen soll schätzen, wie viele Socken, Unterhosen oder Barbie-Puppen in der kommenden Saison voraussichtlich bestellt werden. Zu hohe Vorräte, wenn die Nachfrage überschätzt wird, oder Verzögerungen, wenn keine Ware lieferbar ist, sind ein erheblicher Kostenblock für jeden Versandhändler. Aus Millionen von Bestelldaten der Vergangenheit und Einkaufslisten der Otto-Einkäufer sollen Feindt und sein Team zu-

Physicists modernise sales management ...

The logo for Otto Group, featuring the words "otto group" in a bold, red, lowercase sans-serif font. A thin red curved line starts from the left and sweeps upwards to the right, ending under the "o" in "group".

### Press release

Hamburg, 19 May 2008

#### **Otto Group establishes Joint Venture with Phi-T and revolutionises Item-Sales-Forecasts**

The Otto Group, Hamburg, the largest home-shopping retail group in the world and Physics Information Technologies GmbH, Phi-T, Karlsruhe, have founded the joint venture Phi-T products & services. Hans-Otto Schrader, Chairman of the Executive Board and Chief Executive Officer (CEO) Otto Group and Professor Michael Feindt, Co-founder of Phi-T, announced this today at a press conference in Hamburg. The aim of the co-operation of Otto and Phi-T is the use, exclusive to the sector, of the NeuroBayes Technology of Prof. Feindt, based on artificial neuronal networks to optimize Item-Sales-Forecasts within the Otto Group. Through the co-operation with Phi-T Otto has access to the most up-to-date scientific insights and methods in the area of Data Mining, ensuring itself a significant competitive advantage.

# Prognosis of individual health costs

Pilot project for a large private health insurance

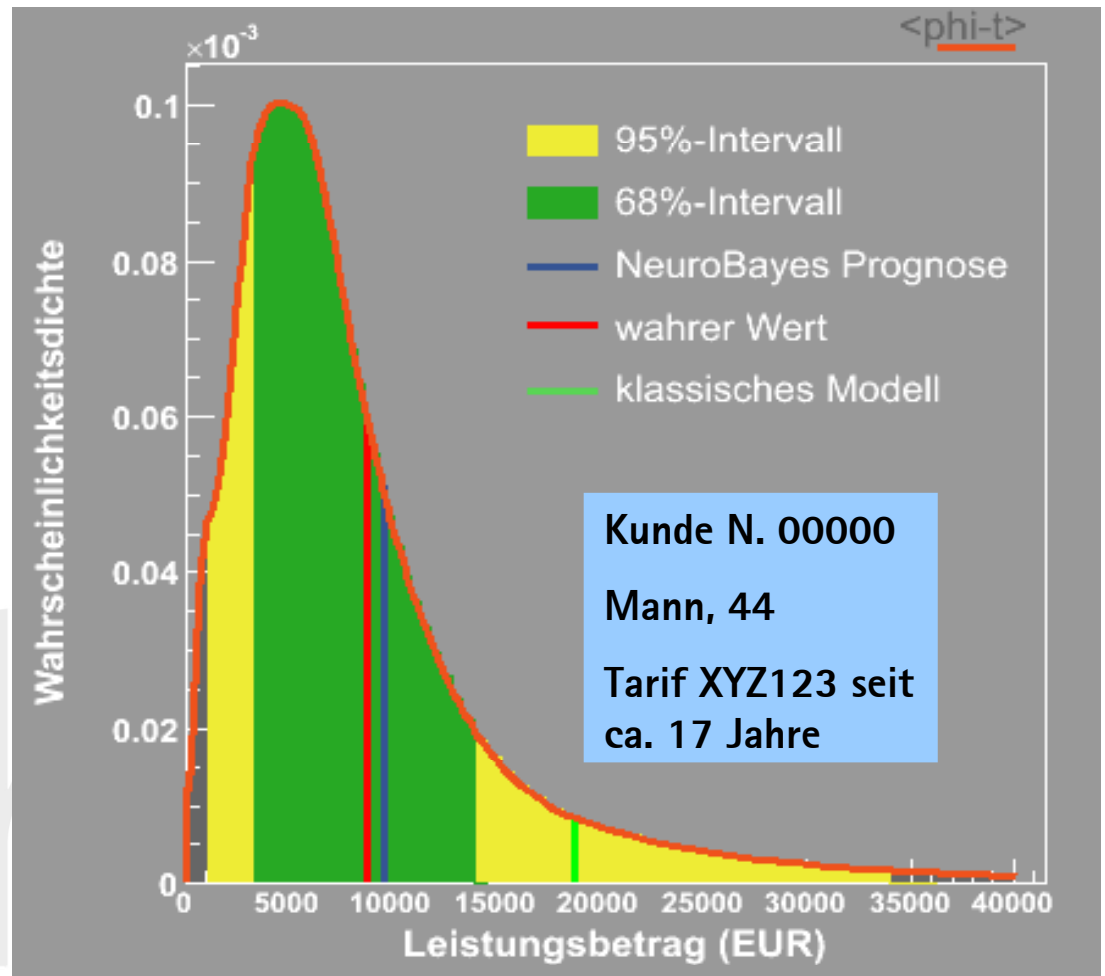
Prognosis of costs in following year for each person insured with confidence intervals

4 years of training, test on following year

Results:

Probability density for each customer/tarif combination

**Very good test results!**



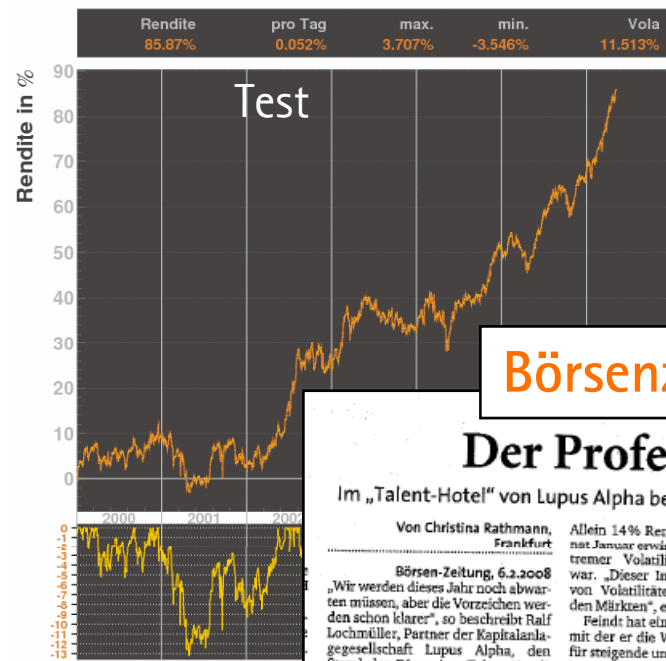
**Has potential for a real and objective cost reduction in health management**

VDI-Nachrichten, 9.3.2007

# Prognosis of financial markets

NeuroBayes® based risk averse market neutral funds for institutional investors

Lupus Alpha NeuroBayes® Short Term Trading Fonds (up to now 20 Mio €, aim: 250 Mio € end of 2008)



Börsenzeitung, 6.2.2008

## Der Professor schafft 26 Prozent

Im „Talent-Hotel“ von Lupus Alpha beginnt die Trennung von Spreu und Weizen – Zwei Ansätze im Minus

Von Christina Rathmann, Frankfurt  
**Börsen-Zeitung, 6.2.2008**  
 „Wir werden dieses Jahr noch abwarten müssen, aber die Vorzeichen werden schon klarer“, so beschreibt Ralf Lochmüller, Partner der Kapitalanlagegesellschaft Lupus Alpha, den Stand der Dinge im „Talent-Hotel“  
 Allein 14% Rendite habe er im Monat Januar erwirtschaftet, der von extremer Volatilität gekennzeichnet war. „Dieser Investmentansatz lebt von Volatilitäten und divergierenden Märkten“, erläutert Lochmüller.  
 Feindt hat eine Formel entwickelt, mit der er die Wahrscheinlichkeiten für steigende und fallende Kurse vorhersagt. Auf Basis der Prognosen im gegenwärtigen Marktumfeld sei nicht geplant, gleich einen Publikumsfonds dafür aufzuliegen. Die beim Start des Talent-Hotels avisierte Testphase von 18 bis 24 Monaten solle eingehalten werden.  
 An zweiter Stelle liegt derzeit Gerd Henning Beck, der ein Rohstoffportfolio verwaltet. Dabei kommen mehrere Modelle zum Einsatz. Er

## Den Elementarteilchen des Börsenparketts auf der Spur

VDI nachrichten, Düsseldorf, 9. 3. 07, ps –  
 Heute wissen, wie morgen die Aktien stehen? Was klingt wie der heimliche Traum eines jeden Börsianers, ist durchaus möglich. Davon ist zumindest Michael Feindt, Elementarteilchen-Physiker an der Universität Karlsruhe, überzeugt.



Hektik auf dem Parkett: Wie in Sao Paulo reagierten Händler weltweit mit Aktienverkäufen auf den Kursrutsch in China. Entwicklungen, die sich künftig vorhersagen lassen? Foto: opa

Das Besondere ist, dass es in der Physik wie an der Börse neben einer überwältigenden Vielzahl zufälliger Ereignisse auch winzige Splitter statistisch signifikanter Zusammenhänge gibt. Es kommt darauf an diese zu ermitteln“, so Michael Feindt. Und damit kennt sich der Physiker aus. Nach 20 Jahren Erfahrung in der statistischen Datenanalyse und Forschungstätigkeiten am Deutschen Elektronen-Synchrotron (DESY) in Hamburg und am European Laboratory for Particle Physics (CERN) in der Schweiz entwickelte er im 2001 das Programm „NeuroBayes“.

Aber wie kommt man vom Labor aufs Börsenparkett? „Als in den späten Neunzigern die Börsenblase platzte, habe auch ich einiges Geld verloren“, erinnert sich der Börsenphysiker. „Und da kam mir die Idee zu prüfen, ob sich nicht mit Hilfe unserer Methode auch Aussagen über die Wahrscheinlichkeit von Entwicklungen in anderen Bereichen machen lassen.“

Christian Haag, Projektmanager der Phi-T, als einer der ersten „Hotel-Bewohner“ nun beweisen, dass Feindts Modell auch an den Aktienmärkten funktioniert.

Der Clou: Die Software findet in riesigen Datenmengen nur statistisch signifikante Strukturen und Zusammenhänge, die auch für Zukunftsprognosen interessant sind, und das so gut wie kaum ein anderes Programm: „In diesem Bereich haben wir heute sicherlich eine der weltweit besten „Technologien“, erklärt Feindt stolz.

Doch gerade in diesem Bereich sind die Vorhersagen ziemlich schwer: „Es gibt nur eine sehr kleine vorhersagbare Komponente in einem, im Wesentlichen absolut zufälligen Markt“, weiß Feindt. Aber: Diese kleine Komponente gibt es und sie liegt in einem Bereich, den Börsenexperten kennen: „Gier und



Physiker Feindt: „Splitter statistisch signifikanter Zusammenhänge.“ Foto: Feindt

Der erste Auftraggeber war eine Versicherung. „Mit Hilfe von NeuroBayes und den historischen Daten über Kunden und Unfälle konnten wir ein optimiertes Risikomodell für Kfz-Versicherungen entwerfen“, erinnert sich Feindt. Es folgten Drogerketten und Bausparbanken. Aktuell ermitteln die Zusammenhänge von Verkaufszahlen und Kataloggestaltung für ein großes deutsches Versandhaus.

Panik! Das sind die zwei wesentlichen Komponenten, die an der Börse zu vorhersagbaren Entwicklungen führen“, so der Physiker.

Einen Artikel über die Erfolge von Michael Feindt lasen die Macher der Frankfurter Fondsgesellschaft Lupus alpha hat im Oktober 2006 das erste „Talent-Hotel“ eröffnet. Die Einrichtung soll

Für die Experten von Lupus alpha ist die Idee aus verschiedenen Gründen interessant: „Zunächst einmal ist es Gold wert, wenn sich jemand gut in der Analyse historischer Daten auskennt und den Möglichkeiten, daraus Prognosen abzuleiten“, erläutert Ulf Becker von Lupus alpha.

Die Small Caps-Spezialisten aus Frankfurt eröffneten im vergangenen Herbst das deutschlandweit erste Talent-Hotel (siehe Kasten). Gemeinsam mit drei weiteren Newcomern kann

Längerfristig versprechen sich die Frankfurter durch Michael Feindt nachhaltig gute und risiko-adjustierte Renditen. Denn: „Ohne Risiko gibt's keine guten Renditen“, so Ulf Becker, „und genau bei dieser Abwägung hilft die Idee definitiv“. Ob dem wirklich so sein wird, bleibt zunächst offen. Aber nicht lange, da ist sich zumindest Michael Feindt sicher: „Zum Glück ist nicht alles Zufall, jedenfalls nicht an den Kapitalmärkten. Deshalb bin ich vom Erfolg unserer Idee felsenfest überzeugt.“ CHRISTOPH GABLER  
 www.phi-t.de



# The <phi-t> mouse game:

or:

even your “free will” is predictable



[//www.phi-t.de/mousegame](http://www.phi-t.de/mousegame)

<phi-t>  
Physics Information Technologies

## Bindings and Licenses

NeuroBayes<sup>®</sup> is commercial software.

Special rates for public research.

Essentially free for high energy physics research.

License files needed. Separately for expert and teacher.

Please contact Phi-T.

NeuroBayes core code written in Fortran.

Libraries for many platforms (Linux, Windows, ...) available.

Bindings exist for C++, Fortran, java, lisp, etc.

Code generator for easy usage exists for Fortran and C++.

**New:** Interface to root-TMVA available (classification only)

Integration into root planned

# Documentation

## Basics:

M. Feindt, A Neural Bayesian Estimator for Conditional Probability Densities, E-preprint-archive physics 0402093

M. Feindt, U. Kerzel, The NeuroBayes Neural Network Package, NIM A 559(2006) 190

## Web Sites:

[www.phi-t.de](http://www.phi-t.de) (Company web site, German & English)

[www.neurobayes.de](http://www.neurobayes.de) (English site on physics results with NeuroBayes & all diploma and PhD theses using NeuroBayes)

[www-ekp.physik.uni-karlsruhe.de/~feindt](http://www-ekp.physik.uni-karlsruhe.de/~feindt) (some NeuroBayes talks can be found here under -> Forschung)



At DELPHI I have learned how to extract information and learn **from data**.

With a physicist's pragmatic analytical and quantitative thinking, data-analysis and programming expertise and NeuroBayes<sup>®</sup>, we have achieved a lot in physics and with Phi-T in very different areas of industry and economics.

Physics Information Technologies

**It is not physics, but the understanding,  
forecasting and steering of complex systems,  
with the tools of a physicist.  
It is (after many years of being a physicist at least)  
equally fascinating.**

**What is absolutely great: our influence is larger!**

**First years were hard and risky: you have to sell!  
There is no public money just coming in ...  
But with many very successful projects we now  
have many contacts to high-ranked people  
convinced of our capabilities.**

In general:

HEP community very effective and well-trained. Standard and motivation much higher than in industry.

Phi-T, NeuroBayes<sup>®</sup> and its industrial applications are a real spin-off from High Energy Physics with its origin in DELPHI.

**btw: There is much to do!**

**Fascinating and challenging projects and permanent positions available. Phi-T will grow!**

Physics Information Technologies