# **Neuroinformatics Lecture (L2)**

Prof. Dr. Gordon Pipa

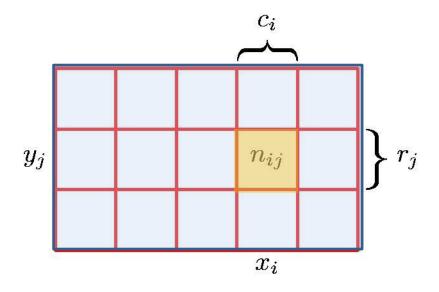
Institute of Cognitive Science University of Osnabrück



Classes of events x and y i.e. :

Classes x: Color: Red or Blue

Classes y: Fruit: Apple or Orange



$$N = \sum_{i,j} n_{i,j} \quad c_i = \sum_j n_{i,j}$$

#### Marginal Probability

(Probability of X=x irrespective of class y)

$$p(X = x_i) = \frac{c_i}{N}.$$

### **Conditional Probability**

(Probability of Y=y given class X=x)

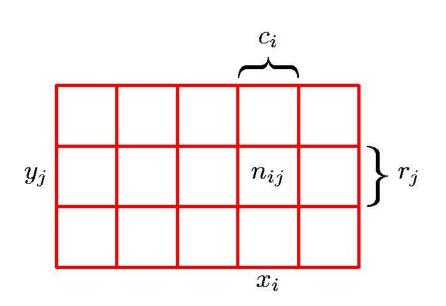
$$p(Y = y_j | X = x_i) = \frac{n_{ij}}{c_i}$$

### Joint Probability

(Probability of X=x and Y=y)

$$p(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$$





Sum Rule  $p(X = x_i) = \frac{c_i}{N} = \frac{1}{N} \sum_{j=1}^{L} n_{ij}$   $= \sum_{i=1}^{L} p(X = x_i, Y = y_j)$ 

joint probability

#### **Product Rule**

$$p(X = x_i, Y = y_j) = \frac{n_{ij}}{N} = \frac{n_{ij}}{c_i} \cdot \frac{c_i}{N}$$

$$= p(Y = y_j | X = x_i) p(X = x_i)$$

joint probability

conditional probability

marginal probability



$$p(X) = \sum_{Y} p(X, Y)$$

$$p(X,Y) = p(Y|X)p(X)$$

#### Thomas Bayes (1701–1761)

Thomas Bayes was born in Tunbridge Wells and was a clergyman as well as an amateur scientist and a mathematician. He studied logic and theology at Edinburgh University and was elected Fellow of the Royal Society in 1742. During the 18th century, issues regarding probability arose in connection with gambling and with the new concept of insurance. One particularly important problem concerned so-called inverse probability. A solution was proposed by Thomas Bayes in his paper 'Essay towards solving a problem in the doctrine of chances', which was published in 1764, some three years after his death, in the Philosophical Transactions of the Royal Society. In fact, Bayes only formulated his theory for the case of a uniform prior, and it was Pierre-Simon Laplace who independently rediscovered the theory in general form and who demonstrated its broad applicability.





$$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)}$$

$$p(X) = \sum_{Y} p(X|Y)p(Y)$$

posterior ∝ likelihood × prior

using 
$$p(X) = \sum_{Y} p(X,Y) \qquad p(X,Y) = p(Y|X)p(X)$$

$$p(X) = \sum_{Y} p(X,Y) = \sum_{Y} p(X|Y)p(Y)$$

- We look for marginal prob. of X=apple
- We know marginal prob. of the box color p(Y)
- We can compute both the joint and conditional prob form the table.

	App.	Org.	Lim.		
r	3	4	3		
b	1	1	0		
g	3	3	4		

- We look for marginal prob. of X=apple
- We know marginal prob. of the box color p(Y)
- We can compute both the joint and conditional prob form the table.

$$p(X = apple) = \sum_{y} p(X = apple, Y)$$
$$p(X = apple) = \sum_{y} p(X = apple | Y)p(Y)$$

	App.	Org.	Lim.		
r	3	4	3		
b	1	1	0		
g	3	3	4		

$$p(X = apple) = \sum_{y} p(X = apple|Y)p(Y)$$

		App.	Org.	Lim.
p(X = apple   Y = red) = 3/10	r	3	4	3
p(X = apple   Y = green) = 3/10	b	1	1	0
p(X = apple   Y = blue) = 1/2	g	3	3	4

$$p(X = apple) = \sum_{y} p(X = apple|Y)p(Y)$$

$$p(X = apple|Y = red) = 3/10$$

$$p(X = apple|Y = green) = 3/10$$

$$p(X = apple|Y = blue) = 1/2$$

$$p(X = apple|Y = green) = 3/10$$

Suppose that we have three colored boxes r (red), b (blue), and g (green).

Box r contains 3 apples, 4 oranges, and 3 limes, box b contains 1 apple, 1 orange, and 0 limes, and box g contains 3 apples, 3 oranges, and 4 limes.

If a box is chosen at random with probabilities p(r) = 0.2, p(b) = 0.2, p(g) = 0.6, and a piece of fruit is removed from the box (with equal probability of selecting any of the items in the box), then **what is the probability of selecting an apple?** 

You may want to use this:

$$p(X) = \sum_{Y} p(X, Y)$$

$$p(X,Y) = p(Y|X)p(X)$$

$$p(X = apple) = \sum_{y} p(X = apple|Y)p(Y)$$

$$p(X = apple) = p(X = apple|Y = red)p(Y = red) +$$
  
 $p(X = apple|Y = blue)p(Y = blue) +$   
 $p(X = apple|Y = green)p(Y = green)$ 

$$= \frac{3}{10} * 0.2 + \frac{3}{10} * 0.6 + \frac{1}{2} * 0.2$$
$$= \frac{6}{100} + \frac{18}{100} + \frac{10}{100} = \frac{34}{100}$$

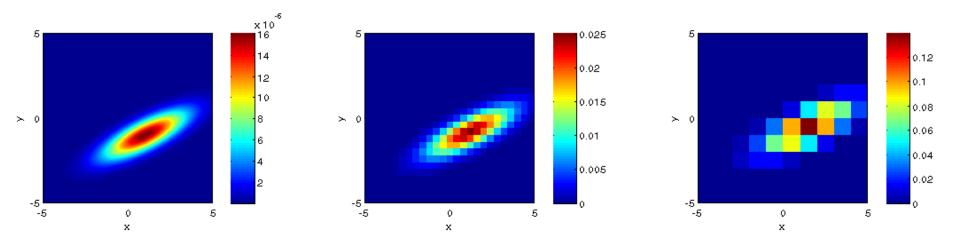
$$p(X = apple | Y = red) = 3/10$$
  
 $p(X = apple | Y = green) = 3/10$   
 $p(X = apple | Y = blue) = 1/2$ 

# **Neuroinformatics Lecture**

# **Topics:**

Continuous and Discrete random variables

# Probability Densities and Probability of Discrete Events



We introduced probabilities for discrete events.

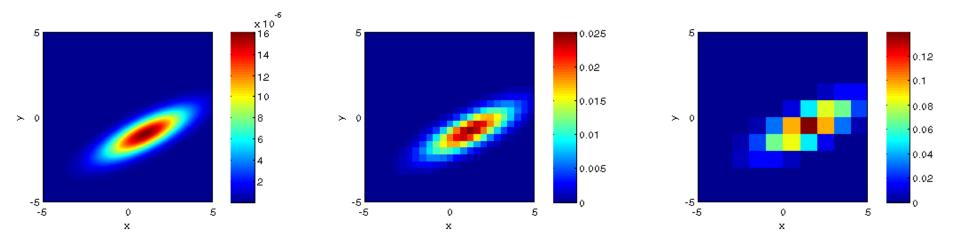
Making the bins arbitrarily small leads to a probability density

$$p(x) \geqslant 0$$

$$\int_{-\infty}^{\infty} p(x) \, \mathrm{d}x = 1$$

See example source code Prob\_Density.m to reproduce the figures

## Probability Densities and Probability of Discrete Events



This implies that, having a probability density, we need to define an interval to define a probability.

#### **Probability**

$$p(x \in (a,b)) = \int_a^b p(x) \, \mathrm{d}x$$

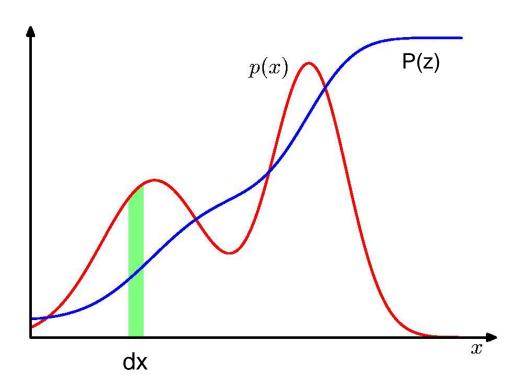
#### **Definitions:**

#### A Discrete Random Variable:

A random variable X is called discrete of the range of X is finite or countably infinite.

#### A Continuous Random Variable:

A random variable X is called continious of the range of X is uncountably infinite.



p(x) is a density

P(z) the CDF

#### PDF (probability density function)

$$p(x) \geqslant 0$$

$$\int_{-\infty}^{\infty} p(x) \, \mathrm{d}x = 1$$

#### **Probability**

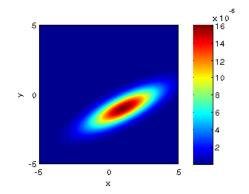
$$P(x \in (a,b) = \int_{a}^{b} p(x)dx$$

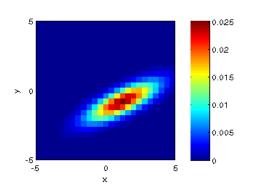
#### **CDF** (cumulative density function)

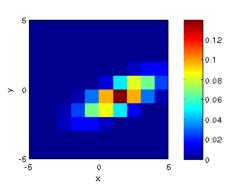
$$P(z) = \int_{-\infty}^{z} p(x) \, \mathrm{d}x$$

What is the difference between a continuous and discrete random variable.

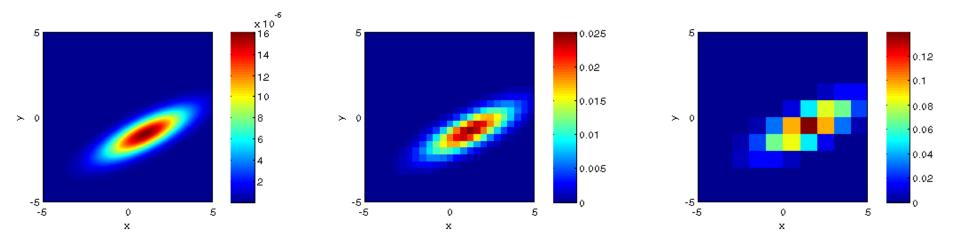
 How can we compute the probability that a continuous random variable takes a value in the range between a and b given that you know its probability density







## Probability Densities and Probability of Discrete Events



This implies that, having a probability density, we need to define an interval to define a probability.

#### **Probability**

$$p(x \in (a,b)) = \int_a^b p(x) \, \mathrm{d}x$$

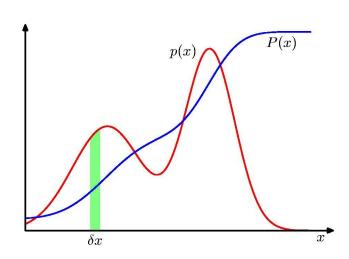
#### A Discrete Random Variable:

A random variable X is called discrete of the range of X is finite or countably infinite.

#### A Continuous Random Variable:

A random variable X is called continious of the range of X is uncountably infinite.

**Probability:** 
$$P(x \in (a, b) = \int_a^b p(x) dx$$



#### **Example HIV Testing:**

#### Suppose:

- A blood test gives a false negative result with p=0.1
   (false negative: negative test if patient is HIV infected)
   (→ test power p=0.9 to detect HIV if patient has HIV)
   p(negative test|has HIV) = 0.1 and p(positive test|has HIV) = 0.9
- A blood test gives a false positive result with p=0.1
   (false positive: positive test if patient is not HIV infected)
   p(positive test|has no HIV)=0.1
- HIV infection is rare: p(has HIV)=0.006

What is the probability that a patient has HIV in the case that he receives a positive test result? p(has HIV | positive test)

- A blood test gives has a test power
- False positive
- HIV infection is rare:

 $p(positive\ test|has\ HIV) = 0.9$   $p(positive\ test|has\ no\ HIV)=0.1$  $p(has\ HIV)=0.006$ 

What is the probability that a patient has HIV in the case that he receives a positive test result? p(has HIV | positive test)

$$p(has\ HIV\ |positive\ test) = \frac{p(positive\ test|has\ HIV)p(has\ HIV)}{p(positive\ test)}$$

#### We used Bayes

$$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)}$$

- A blood test gives has a test power
- False positive
- HIV infection is rare:

 $p(positive\ test|has\ HIV) = 0.9$   $p(positive\ test|has\ no\ HIV)=0.1$  $p(has\ HIV)=0.006$ 

What is the probability that a patient has HIV in the case that he receives a positive test result? p(has HIV | positive test)

$$p(has\ HIV\ |positive\ test) = \frac{p(positive\ test|has\ HIV)p(has\ HIV)}{p(positive\ test)}$$

We use sum rule (,we marginalize'): Add all p for a positive test

$$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)} \qquad p(X) = \sum_{Y} p(X,Y) = \sum_{Y} p(X|Y)p(Y)$$

- A blood test gives has a test power
- False positive
- HIV infection is rare:

 $p(positive\ test|has\ HIV) = 0.9$   $p(positive\ test|has\ no\ HIV)=0.1$  $p(has\ HIV)=0.006$ 

What is the probability that a patient has HIV in the case that he receives a positive test result? p(has HIV | positive test)

$$p(has\ HIV\ |positive\ test) = \frac{p(positive\ test|has\ HIV)p(has\ HIV)}{p(positive\ test)}$$

We use sum rule (,we marginalize'): Add all p for a positive test

$$\frac{p(positive\ test)}{p(positive\ test)} = p(positive\ HIV) + p(positive\ no\ HIV) \qquad \text{sum rule}$$

$$= p(positive\ HIV)p(HIV) + p(positive\ no\ HIV)p(no\ HIV)$$

$$p(positive\ no\ HIV) + p(positive\ no\ HIV)p(no\ HIV)$$

$$p(positive\ no\ HIV) + p(positive\ no\ HIV)p(no\ HIV)$$

- A blood test gives has a test power
- False positive
- HIV infection is rare:

$$p(positive\ test|has\ HIV) = 0.9$$
  
 $p(positive\ test|has\ no\ HIV)=0.1$   
 $p(has\ HIV)=0.006$ 

What is the probability that a patient has HIV in the case that he receives a positive test result? p(has HIV | positive test)

$$p(has\ HIV\ |positive\ test) = \frac{p(positive\ test|has\ HIV)}{p(positive\ test)}$$

$$p(\textit{has HIV} | \textit{positive test}) = \frac{p(\textit{positive}|\textit{HIV})p(\textit{HIV})}{p(\textit{positive}|\textit{HIV})p(\textit{HIV})} + \frac{p(\textit{positive}|\textit{no HIV})p(\textit{no HIV})}{p(\textit{positive}|\textit{HIV})p(\textit{HIV})}$$

$$p(has\ HIV\ |positive\ test) = \frac{(1-0.1)\cdot 0.006}{(0.9)\cdot 0.006 + 0.1\cdot (1-0.006)} = \frac{0.0054}{0.0054 + 0.0994}$$

- A blood test gives a false negative result with p=0.1 → test power 0.9
- A blood test gives a false positive result with p=0.1
- HIV infection is rare: p=0.006

What is the probability that a patient has HIV in the case that he receives a positive test result?

$$p(has\ HIV\ |positive\ test) = \frac{(1-0.1)\cdot 0.006}{(1-0.1)\cdot 0.006 + 0.1\cdot (1-0.006)} = 0.0515$$

The probability that you have HIV when you get a positive test result is p=0.05. (Keep in mind the numbers a not real!)

Consequence: The HIV test alone is not sufficient but an important marker that indicates further medical check ups!



# The 4W question session

What is this it all about?

What can I use it for?

Why should I learn it?

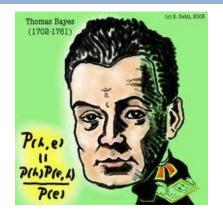
What are potential Bachelor and Master thesis topics?



# <u>The Brain is Bayesian</u>









Review

TRENDS in Cognitive Sciences Vol.10 No.7 July 2006

Full text provided by www.sciencedirect.com

Special Issue: Probabilistic models of cognition

# Bayesian decision theory in sensorimotor control

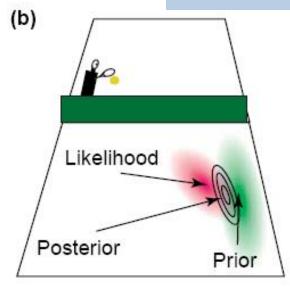
Konrad P. Körding<sup>1</sup> and Daniel M. Wolpert<sup>2</sup>

<sup>1</sup>Brain and Cognitive Sciences, Massachusetts Institute of Technology, Building NE46-4053, Cambridge, Massachusetts, 02139, USA <sup>2</sup>Department of Engineering, University of Cambridge, Trumpington Street, Cambridge, CB2 1PZ, UK

### Paper is uploaded in Studip







Decision theory quantifies how people should choose in the context of a given utility function and some partial knowledge of the world. The expected utility is defined as:

$$E[Utility] \equiv \sum_{\substack{possible \\ outcomes}} p(outcome|action)U(outcome)$$

where p(outcome|action) is the probability of an outcome given an action and U(outcome) is the utility associated with this outcome.

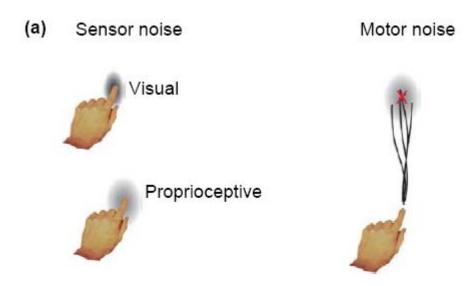
#### Outcome defined by:

- Prior of ball position
- Likelihood of ball position
- Prior of success in making a point dependent on the position

#### Is the brain really Bayesian?

#### **Experiment:**

- A subject needs to point to a target.
- There are three types of noise leading uncertainty



# 4W

#### Box 2. Bayesian statistics

When we have a Gaussian prior distribution p(x) and we have a noisy observation o of the position that leads to a Gaussian likelihood (red curve, Figure I) p(o|x) it is possible to use Bayes rule to calculate the posterior distribution (yellow curve, Figure I; how probable is each value given both the observation and the prior knowledge):

$$p(x|o) = p(o|x)\frac{p(x)}{p(o)}$$

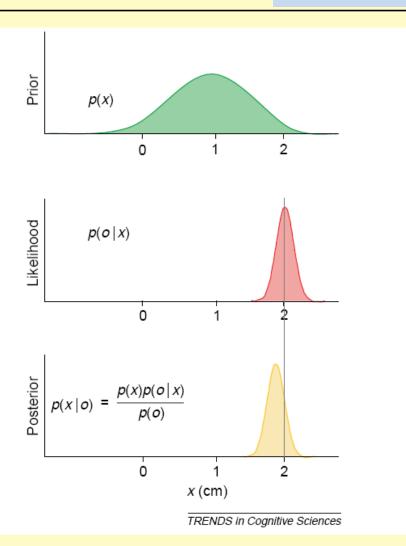
This equation assigns a probability to every possible location. If we assume that the prior distribution p(x) is a symmetric one dimensional Gaussian with variance  $\sigma_p^2$  and mean  $\hat{\mu}$  and that the likelihood p(o|x) is also a symmetric one dimensional Gaussian with variance  $\sigma_o^2$  and mean o, it is possible to compute the posterior that is then also Gaussian in an analytical way. The optimal estimate  $\hat{x}$ , that is the maximum of the posterior is:

$$\hat{x} = \alpha o + (1 - \alpha)\hat{\mu}$$

where

$$\alpha = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_p^2}$$

Moreover we can calculate the width of the posterior as  $\sigma^2 = \alpha \sigma_o$ . The parameter  $\alpha$  is always less than 1. This Bayesian approach leads to a better estimate of possible outcomes than any estimate that is only based on the sensory input.



**Figure I.** Bayesian integration. The green curve represents the prior and red curve represents the likelihood. The yellow curve represents the posterior, the result from combining prior and likelihood.

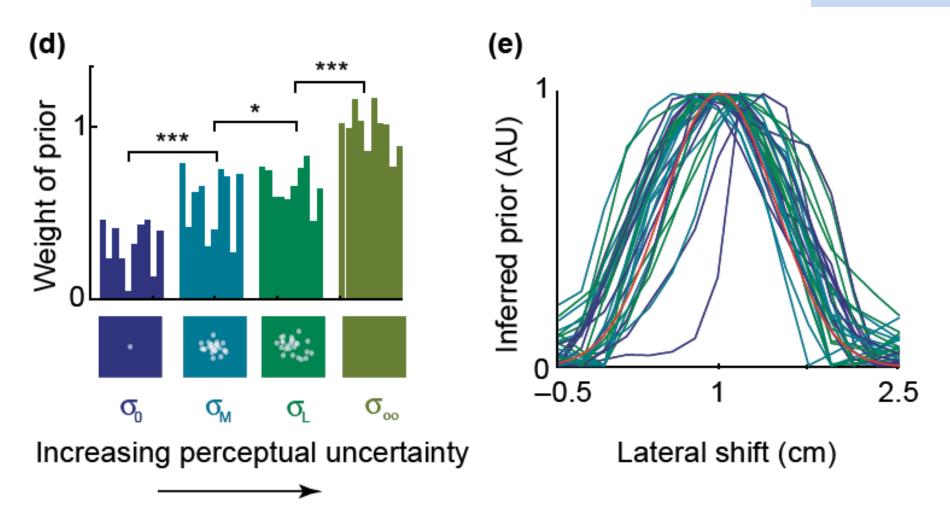
#### Is the brain really Bayesian?

#### **Experiment:**

- A subject needs to point to a target.
- There are three types of noise leading uncertainty

#### Test:

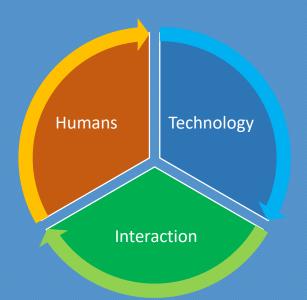
- Can the subject learn the uncertainty of the individual sources to estimate the width of the likelihood?
- Can the subject learn a prior?
- Can it combine the prior and likelihood such that it makes a decision that is maximizing the posterior?



Yes we can, actually for this example it is Baysian optimal!

### **Cognitive Computing**





Symbiotic fusion of the intelligent system, the user, and the expert.

Dialog between machine and human (natural language, intuitive graphics, and gestures)

Neuro-inspired hardware for fault-tolerant new computing devices

The machine is the super assistant that enables the human to make truly intelligent decisions in complex scenarios.



At the core of Cognitive Computing since 15 years

10 Full professors

~ 600 BSc

~ 200 MSc

~ 45 PhD students

# Institute of Cognitive Science and IBM Watson



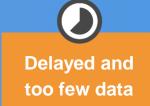


## Why Cognitive Computing?









# **F** flu prediction

A one year project by a core team of three master's students























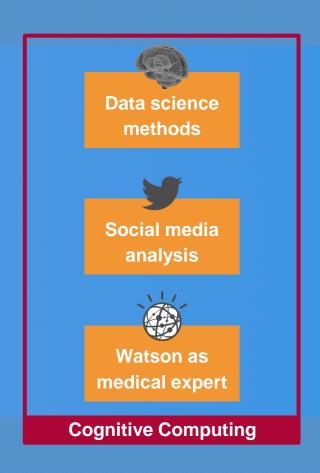


## 😝 flu prediction











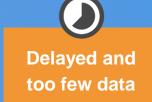
























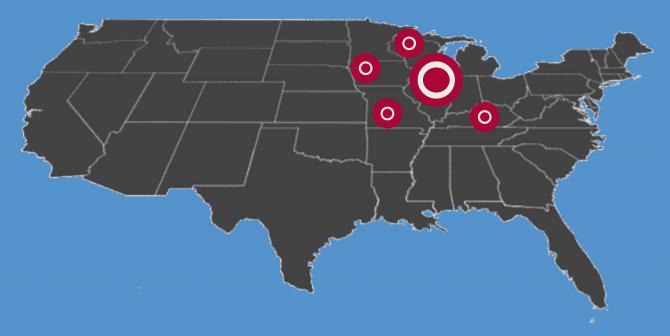
# 🔁 flu prediction



- Disease spreads locally and via transportation hubs
- Weather, vaccination, and seasonal events change spreading



# 🔁 flu prediction



- Disease spreads locally and via transportation hubs
- Weather, vaccination, and seasonal events change spreading



## 🔁 flu prediction



- Disease spreads locally and via transportation hubs
- Weather, vaccination, and seasonal events change spreading



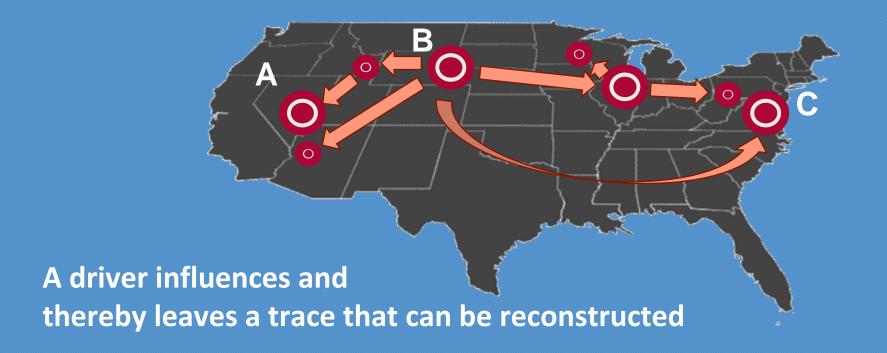




Direction and speed of spread NEEDS to be identified from data



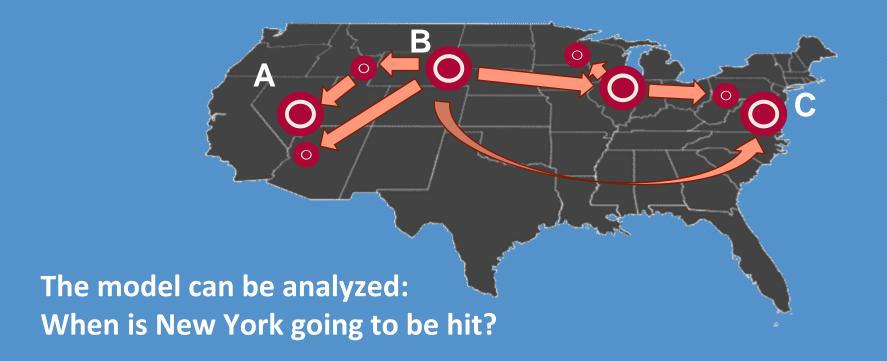




- Schumacher et al. (2015) A Statistical Framework to Infer Delay and Direction of Information ...
- Sugihara et al. (2012) Detecting Causality in Complex Ecosystems







- Schumacher et al. (2015) A Statistical Framework to Infer Delay and Direction of Information ...
- Sugihara et al. (2012) Detecting Causality in Complex Ecosystems





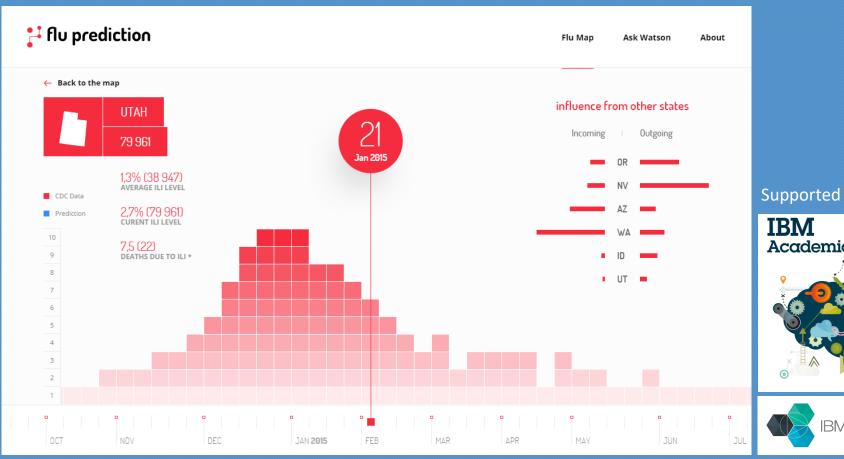


- Schumacher et al. (2015) A Statistical Framework to Infer Delay and Direction of Information ...
- Sugihara et al. (2012) Detecting Causality in Complex Ecosystems

#### IBM Blue Mix & Watson at work



## flu prediction



#### Supported by:





#### IBM Blue Mix & Watson at work







#### Supported by:



IBM Bluemix™





















#### Sample Tweets from 29/09/16





chaos tK ② @WhosChaos · 10 Std.
Really hope I'm not getting the flu ♀️

himself worried



**Kim** @nanosounds · 1. Okt. Anyone had any experience of getting better from **flu**,then getting worse?I was up and about yesterday, but today I'm exhausted and **sick** again herself sick



Rob Sinclair @RSinclairAuthor · 29. Sep.
It's that time of the year again...the school/nursery flu merry-go-round - both boys sick tonight! See you on the other side in April...

familiy is sick



**Halen Sumner** @haysum10 · 29. Sep.
The Centenary **flu** has started making its way around campus. For those who don't wish to die: cover yo mouth, wash yo hands, & shun the **sick** 

friends are sick







Twitter activity (geo tag + tweet)

Twitter gives realtime and anytime available data IBM Insights contains tweets since 2014





Realtime fuzzy social media



Slow but reliable CDC data



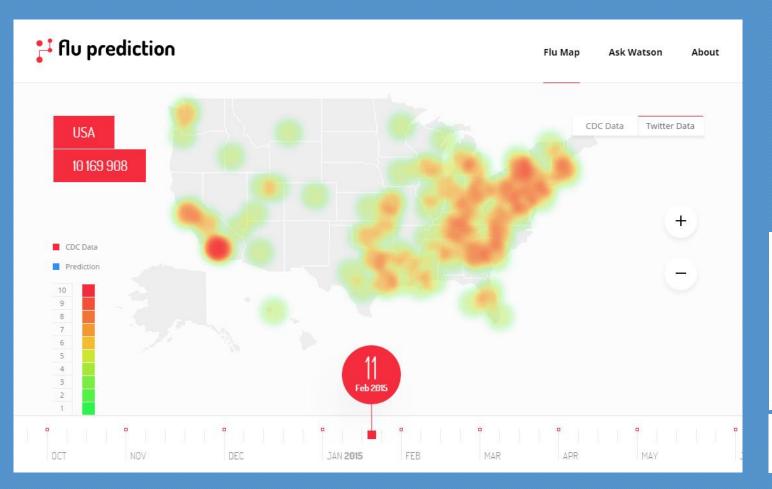
CDC - delayed influenza data

Use the best from both worlds to improve prediction

#### IBM Blue Mix & Watson at Work



## 😝 flu prediction



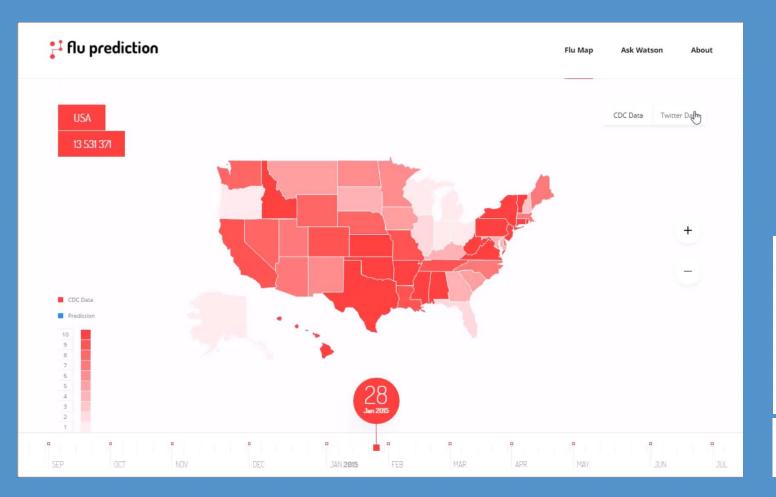
#### Supported by:



#### IBM Blue Mix & Watson at Work



# 🔁 flu prediction

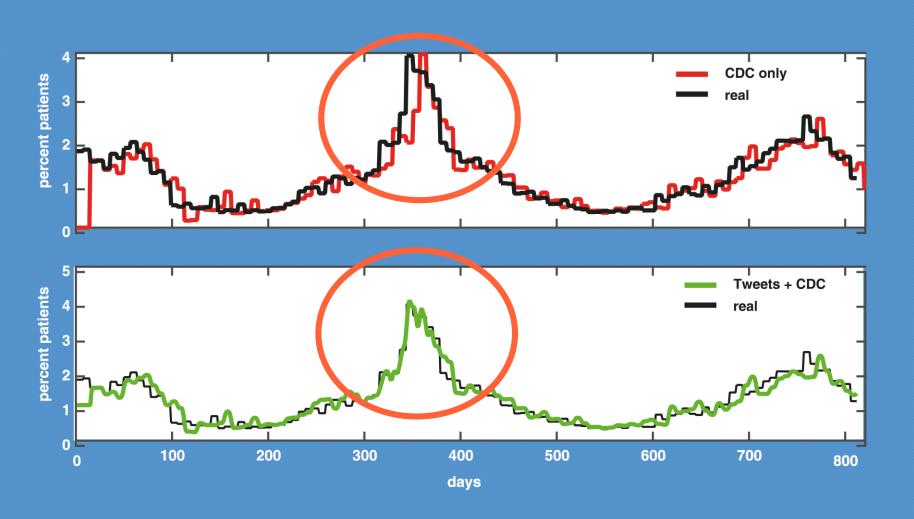


#### Supported by:



## Performance – One Region USA



















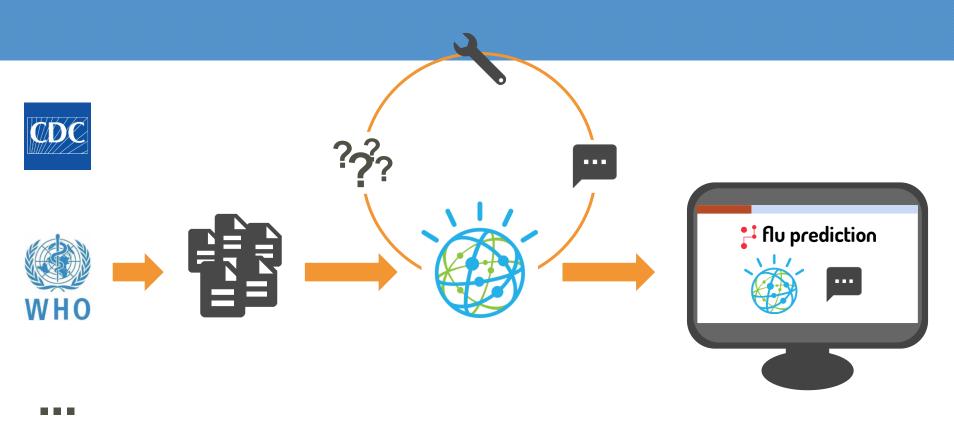






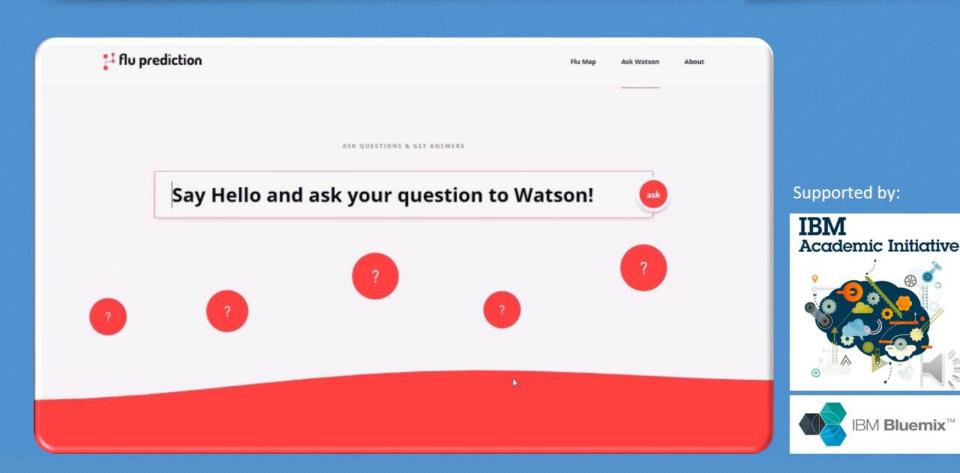


## Flu prediction



## Speak with Watson Flu











- Data science allows identification of very complex causal relations
- Efficient use of BLUE Mix and Watson services



 Combine social media with other conventional data to get the best of both worlds → realtime and reliable



- Use large corpora to identify structure and relationships in your problem
- Use natural language interface for easy to use HCI



# H flu prediction







Prof. Dr. Gordon Pipa University Osnabrück gpipa@uos.de

Prof. Dr. Kühnberger University Osnabrück

Prof. Dr. Dr. Scheller University Hospital Frankfurt

## Study Project - Pepper TA





#### **COGNITIVE COMPUTING HACKATHON 2016**









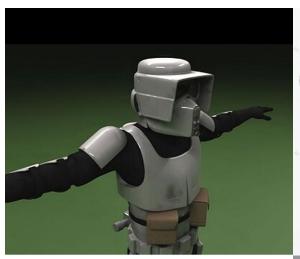
## Our Lab



- HMD with see through option for argumented reality
- Optical and inertia tracking
- Live Characters
- Eye tracking integrated in the HMD

# 4W

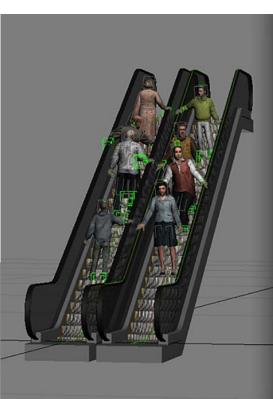
## Free Databases for 3D models: http://custom-3d.turbosquid.com/











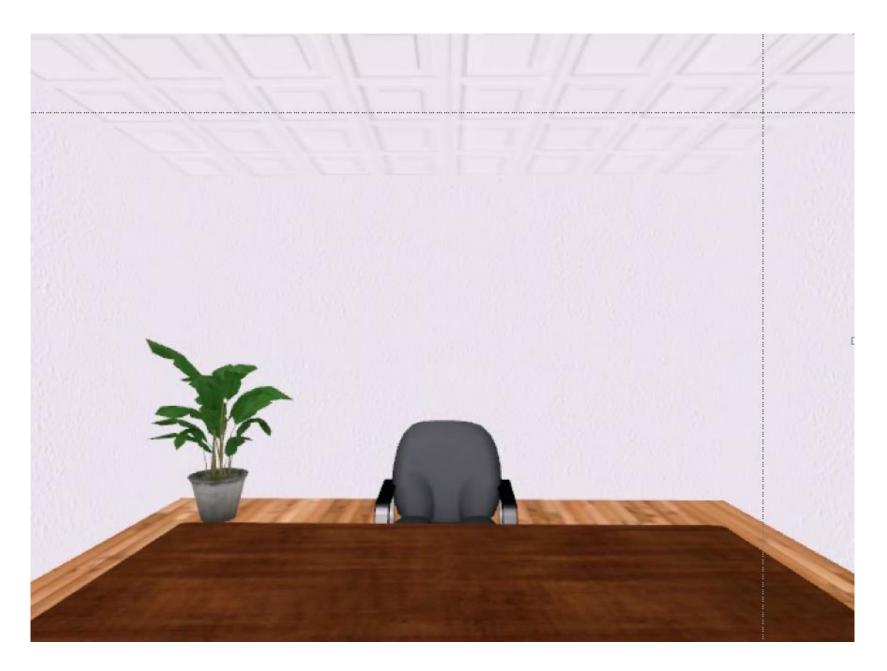
# Virtual Reality Lab

**Example Workflow** 

AG Neuroinformatik

Universität Osnabrück





By Felix Breuniger and Gordon Pipa (The Ultimate Game in VR - neuroeconomics)

