Chances with deep learning in animal sciences?!

June 12th, 2019 – Wageningen Data Science Meet-up

Erwin Mollenhorst, Wageningen Livestock Research

WAGENINGEN
UNIVERSITY & RESEARCH

100years
1918 — 2018

Outline

- Data sources
- Techniques
- Examples
- Conclusions & questions

Sources of Big Data - Machines

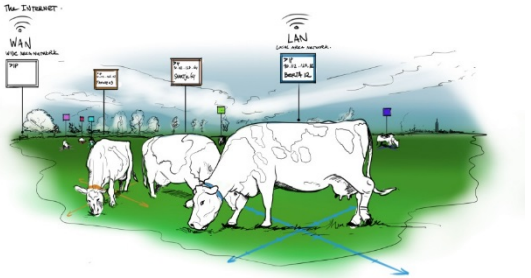
- Tractors
- Tillage equipment
- Milking robot / parlour
- Feed boxes
- .....

Sources of Big Data - Fields

- Soil analysis
- Soil type
- Soil temperature
- Ground water level
- Crop history
- .....

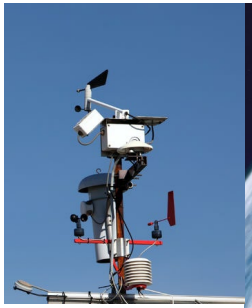
Sources of Big Data - Animals

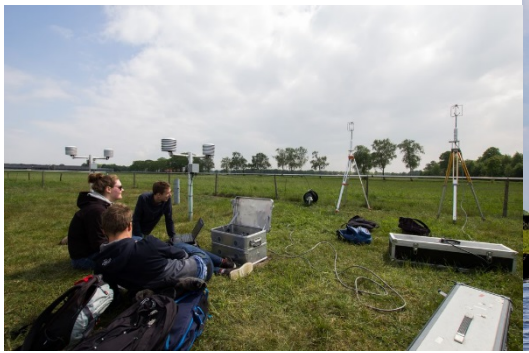
- Genomic data
- Sensors / images
  - ID
  - Behaviour
  - Health
  - Position
  - Smart fencing
- .....



Sources of Big Data - Environment

- Gaseous emissions
  - Methane (CH_4)
  - Ammonium (NH_3)
  - Nitrous oxide (N_2O)
- Ground/surface water
- Weather
- .....



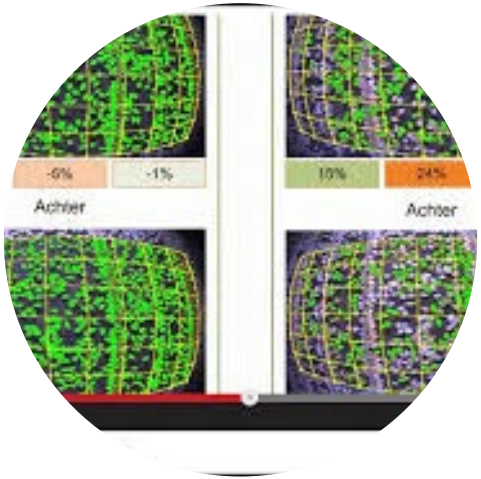


Sources of Big Data – production chain

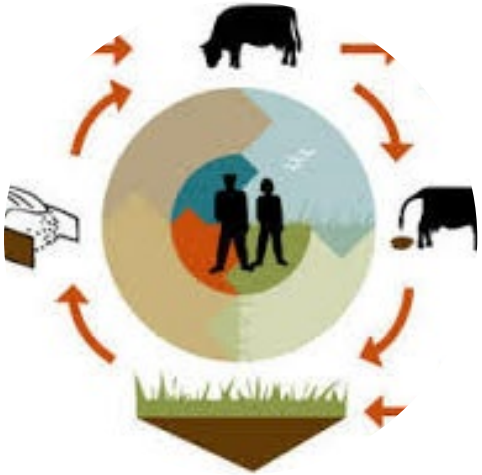
- Slaughter data
- Tracking & tracing
- Farm management program
- Financial accounts
- .....

Big Data & Wageningen Livestock Research

Management tools



Sensor technologies



Food chain

Applications in Big Data projects

- Animal behaviour / tracking of animals
- Broiler/pig production chain
- Pig performance
- Dairy cow’s longevity
- Resilience and efficiency of animal and farms
- Feed intake
- Environmental impact
    - Manure management
    - Emissions from farm or animal

Used techniques

Numerical data

- Ensemble (tree) methods (random forest, GBM)
- Neural networks (extreme learning machine, NN)
- K-nearest neighbour
- Bayesian networks

Images

- Convolutional neural networks

Big Data & Wageningen Livestock Research

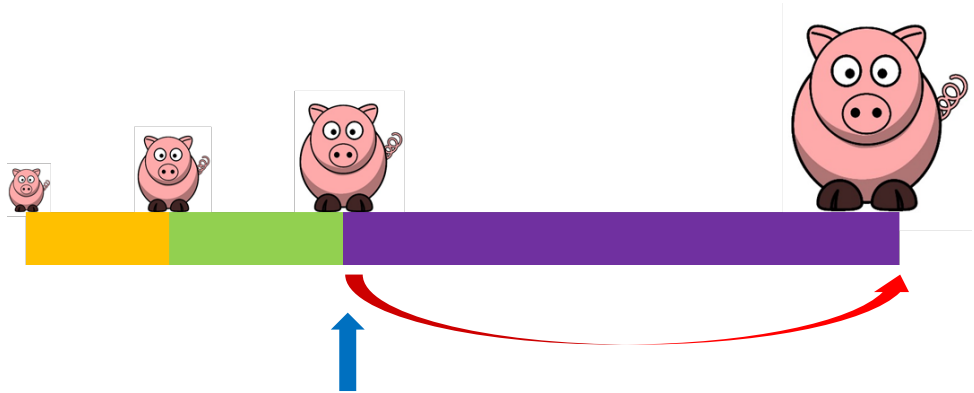
Pig performance

Erwin Mollenhorst, Karel de Greef, Bart Ducro, Ina Hulsegge,

Rita Hoving, Roel Veerkamp, Claudia Kamphuis

Research project – pig management

To predict deviant slaughter pigs
based on routine data available at
the onset of the growing-finishing phase



Dataset from VIC Sterksel

65,208 records of individual pigs

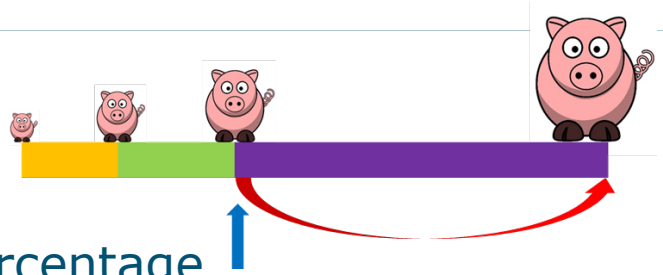
Born between 2004 – 2016

GBM (boosted trees)

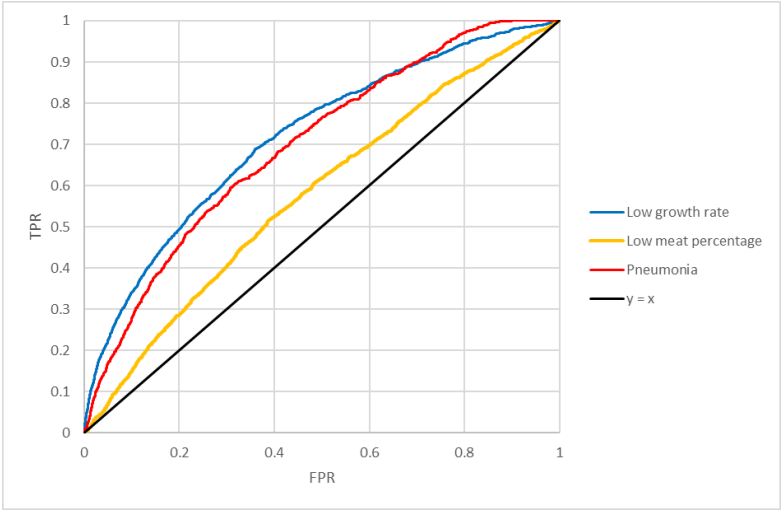
Information on:

- Offspring, litter
- Locations, transfer dates, weights
- Slaughterhouse data

Conclusions



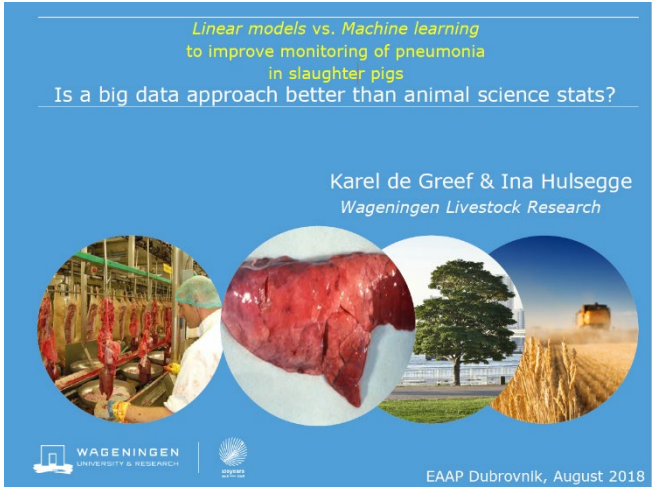
- No reasonable prediction for low meat percentage
- Moderate for pneumonia and low growth rate
- First step towards early warning system



Big Data & Wageningen Livestock Research

Pig slaughter data

Ina Hulsegge & Karel de Greef



Random forest best performing

Results in the ideal dataset (119 farms)

	test	*predict*
Linear Modelling:	r^2	r^2
• Linear regression*	0.32	0.29
ML Methods		
• **Random Forest***	**0.41**	**0.41**
• Gradient Boosting*	0.35	0.34
• GLMnet Lasso*	0.32	0.29
• Extreme Learning Machine	0.22	0.25
• K Nearest Neighbors*	0.18	0.18
• Neural networks	..	

WAGENINGEN UNIVERSITY & RESEARCH

* incl variants (reduced model, transformed or not)
~same order for RMSE and for MAE

5

WAGENINGEN UNIVERSITY & RESEARCH

Conclusions

- RF (ML) is better in predicting on novel data than linear regression
- However, differences depends on disease incidence

- Computers can not take over thinking:

  assessment of the real improvement needed

Big Data & Wageningen Livestock Research

Dairy cow's longevity

Esther van der Heide, Bart Ducro, Roel Veerkamp, et al.

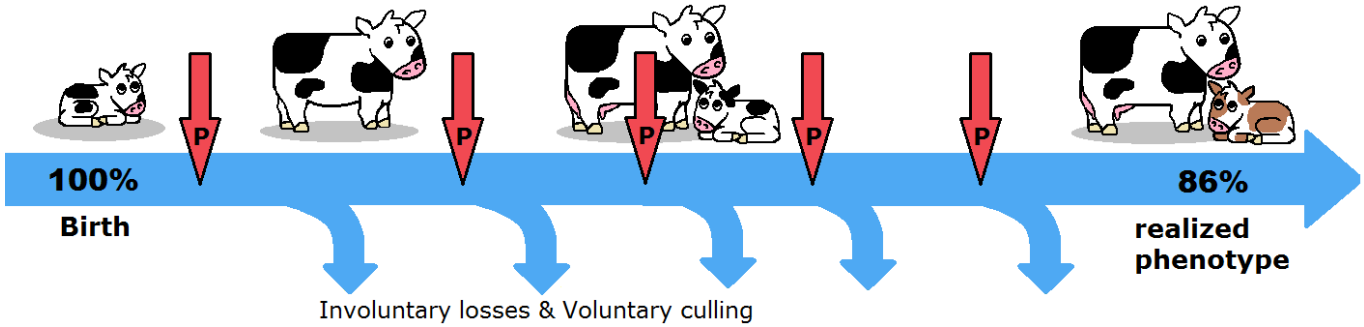
WAGENINGEN UNIVERSITY & RESEARCH

100years 1918 — 2018

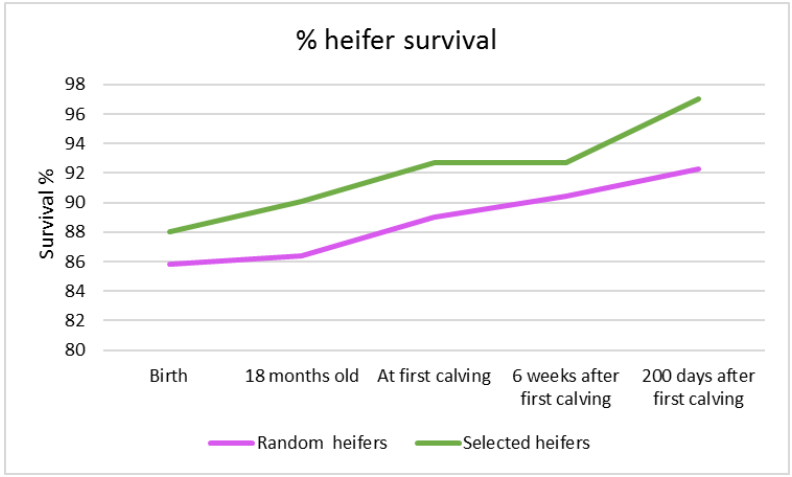
university of groningen

NWO Applied and Engineering Sciences

Dairy cow's longevity



- Important for economics, management and society
- Different machine learning tech.
- 'Informative missingness'
- Different ways of pre-processing
- Neural networks



WAGENINGEN UNIVERSITY & RESEARCH

Big Data & Wageningen Livestock Research

Environmental impact

Manure management

Erwin Mollenhorst, Gerard Migchels, Michel de Haan, Jouke Oenema,

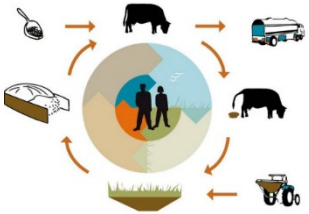
Rita Hoving, Roel Veerkamp, Claudia Kamphuis, et al.

(Be)Mest wijs

step by step towards zero artificial fertilizer

Farm

Annual Nutrient Cycling Assessment (ANCA)



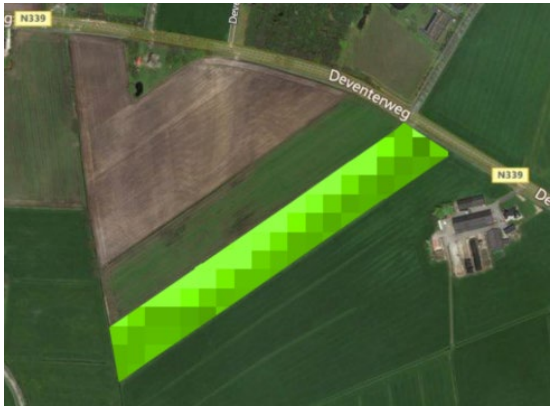
KRINGLOOP WIJZER

Field

Akkerweb

Within field

Precision fertilization



current → short term → (semi) long term

WAGENINGEN UNIVERSITY & RESEARCH

100years 1918 — 2018

First trials

Can we predict future crop yields (= P) based on farm data and open source weather data?

1 farm, 20 years of data, moderate prediction accuracy

Questions:

How to deal with data from different farms, regions, soil types, etc.?

How to utilize different layers of information?

Drone images

GENTORE Task 3.3

Machine Learning

Jappe Franke, Sander Mücher,
Henk Kramer, Ben Loke

Big Data Network meeting, Lunch presentatie, FORUM, 16 Mei 2019

Contact: Jappe Franke, Wageningen Environmental Research (Alterra)
Tel: +31 (0) 317 481672, email: jappe.franke@wur.nl

(M12-M48)

This project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No 727213

Summary & Questions

- Different machine learning techniques are applied in animal science
- Data availability and data integration are often a problem
- Technology is not the silver bullet!

- Can deep learning provide opportunities beyond other techniques?
- What type of case / data set would be suitable for DL?
- How much data is needed for DL?

Chances with deep learning in animal sciences?!

For which challenges in animal sciences could deep learning be a solution?

Erwin.Mollenhorst@wur.nl

WAGENINGEN UNIVERSITY & RESEARCH

100years 1918 — 2018