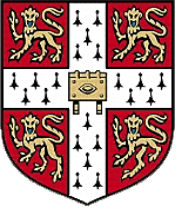


Radu Calin Dimitriu

NEURAL NETWORKS, TRICKS OF THE TRADE

BEFORE STARTING THE TRAINING

EXAMINE THE DATA VERY CLOSELY

AFTER A CLOSE EXAMINATION WE CAN ESTABLISH

- a)

  WE HAVE GOOD DATA

- b)

  WE HAVE BAD DATA

HOW TO DEFINE GOOD DATA?

GOOD DATA ARE PHYSICALLY MEANINGFUL AND RELEVANT TO THE OUTPUT

FOR EXAMPLE

TO MODEL THE STRENGTH OF A STEEL....

- WE MUST CONSIDER THE CHEMICAL COMPOSITION, THE HEAT TREATMENT, GRAIN SIZE...

- IT WOULD BE A BAD IDEA TO INSERT THE MACHINING SPEED

GOOD DATA

- ARE MADE OF AN OPTIMISED NUMBER OF INPUT PARAMETERS
- IT IS NOT RECOMMANDED TO HAVE AN OVER-AMBITIOUS SET OF INPUT PARAMETERS
- BECAUSE THAT WILL LIMIT THE DATA AVAILABLE FOR ANALYSIS

BE REALISTIC

- ALWAYS MAKE A PRAGMATIC APPROACH WHEN MAKING A DATABASE FOR A NEURAL NETWORK

- THIS WILL ENSURE SUFFICIENT DATA AND VARIABLES TO CAPTURE THE COMPLEXITY OF THE PROBLEM

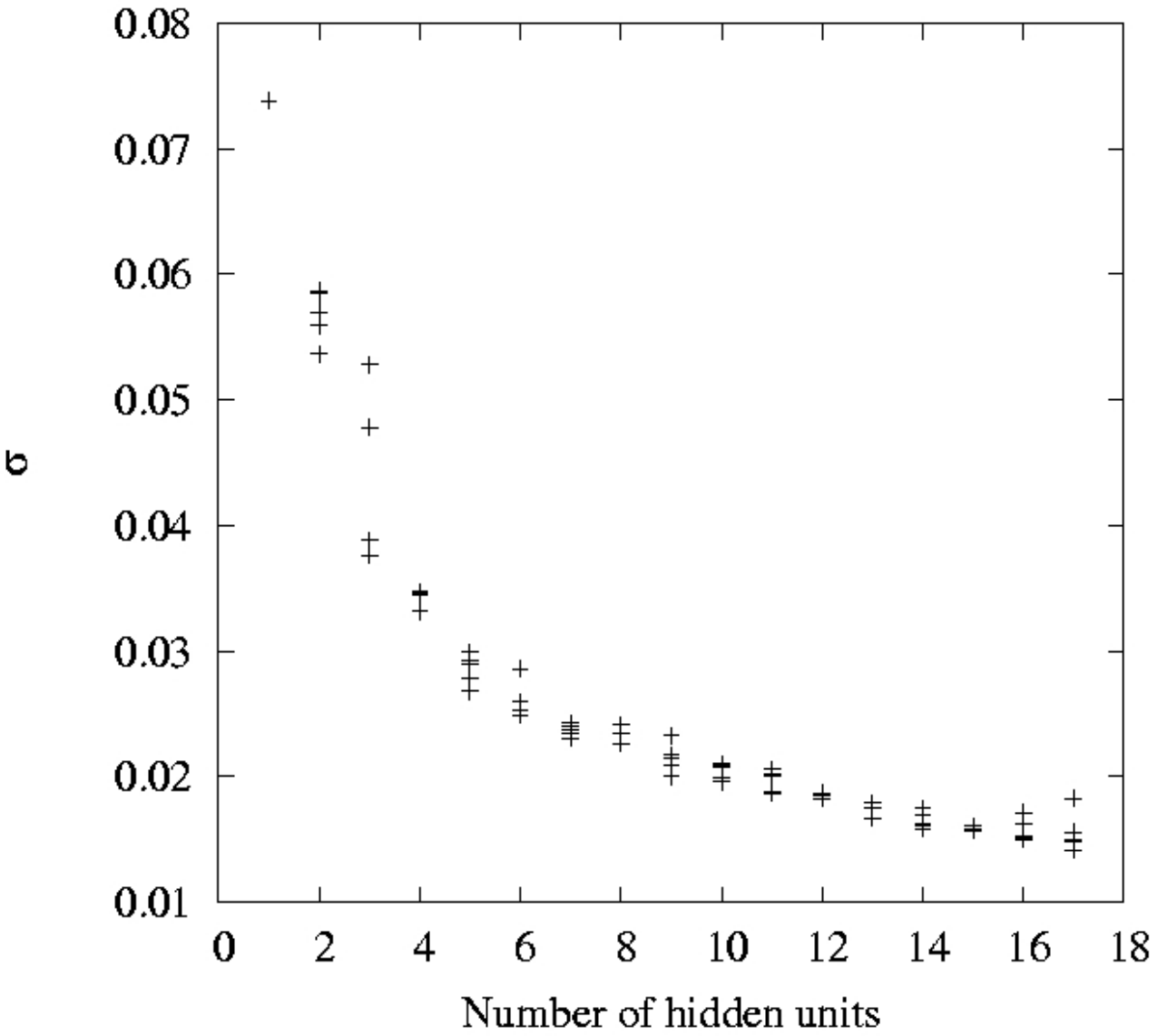
INCLUDE AS MANY MEANINGFUL PARAMETERS AS POSSIBLE

- INCLUDE MEANINGFUL PARAMETERS IN THE DATABASE

- IF YOU DO NOT KNOW HOW THEY ARE RELATED TO THE OUTPUT, THE NEURAL NETWORK MAY NEVERTHELESS CAPTURE THE RELATION

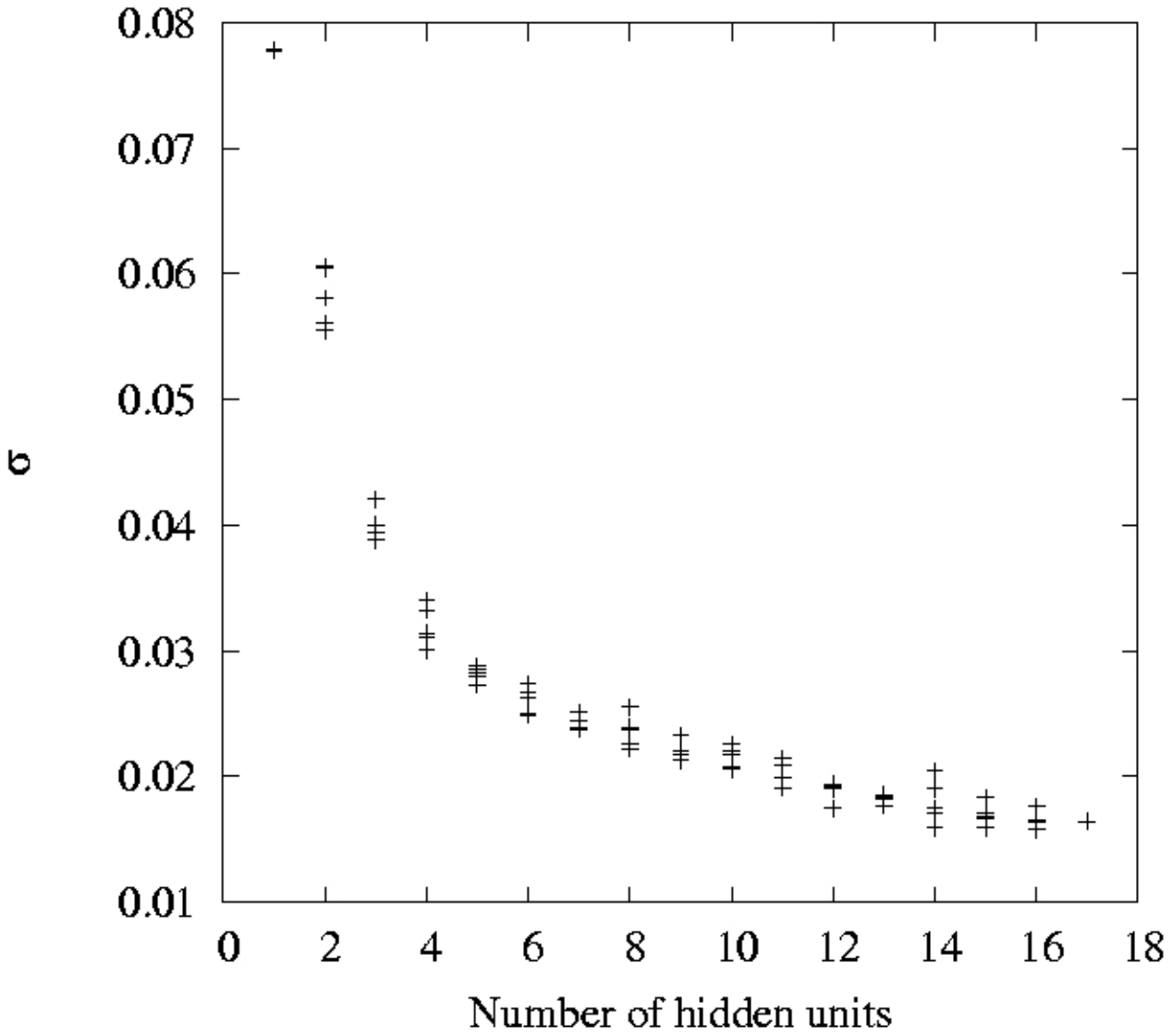
LEVEL OF PERCEIVED NOISE

THE LEVEL OF NOISE FOR A MODEL WITH 26 INPUTS



LEVEL OF PERCEIVED NOISE

THE LEVEL OF NOISE FOR A MODEL WITH 22 INPUTS

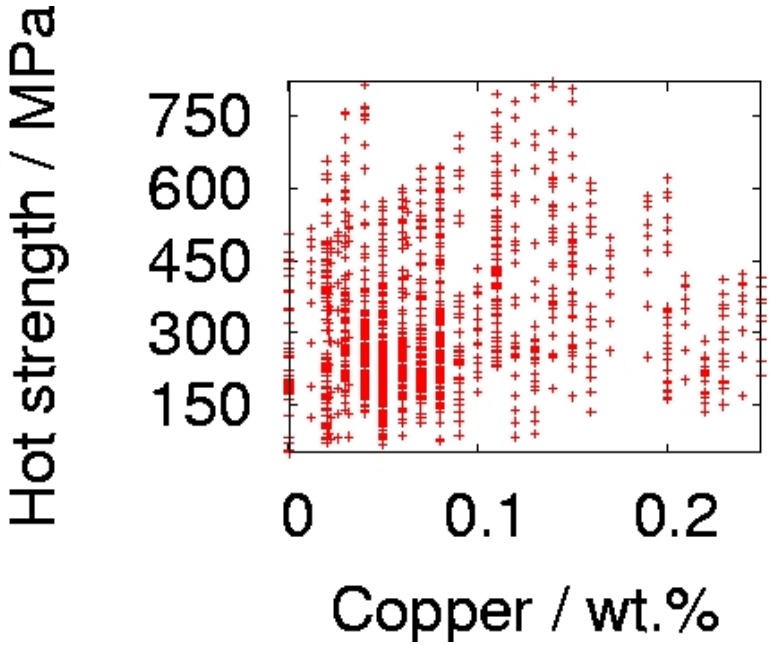


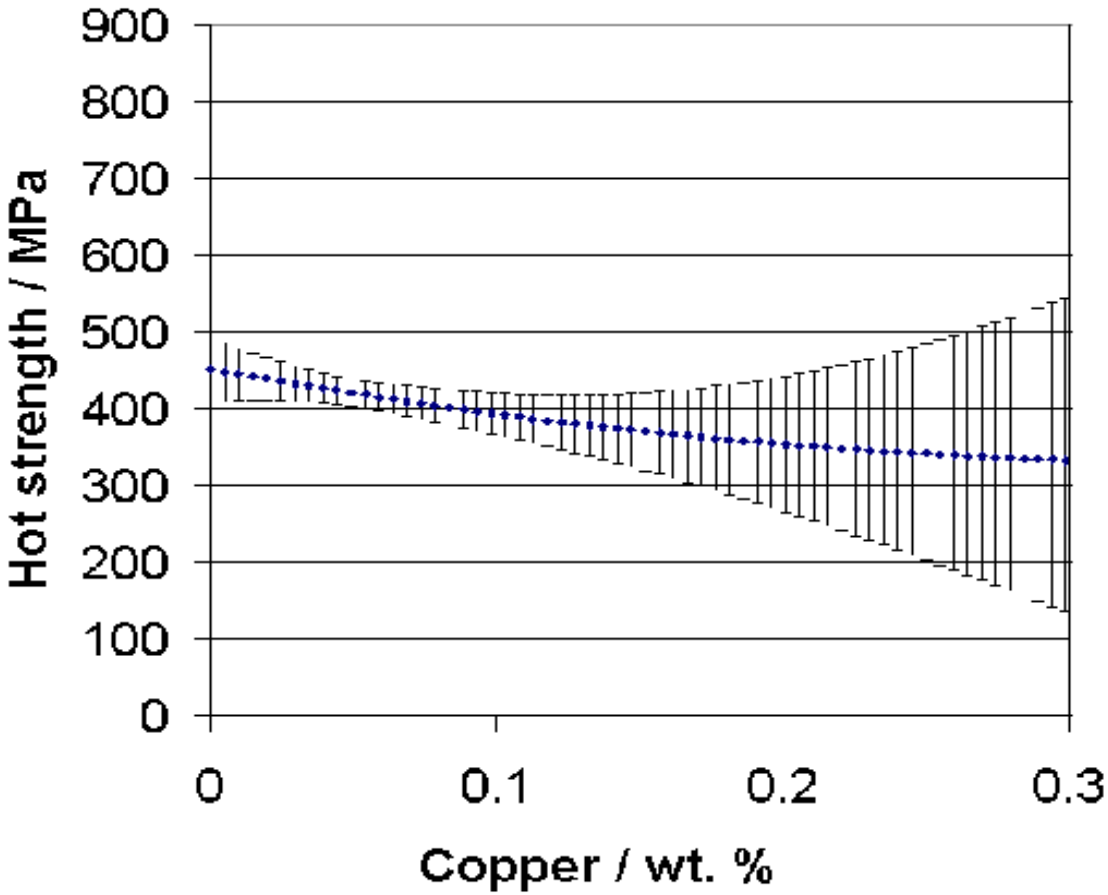
FOR A GOOD DATABASE

- IT IS DESIRABLE TO HAVE A GOOD SPREAD OF DATA OVER THE ENTIRE RANGE

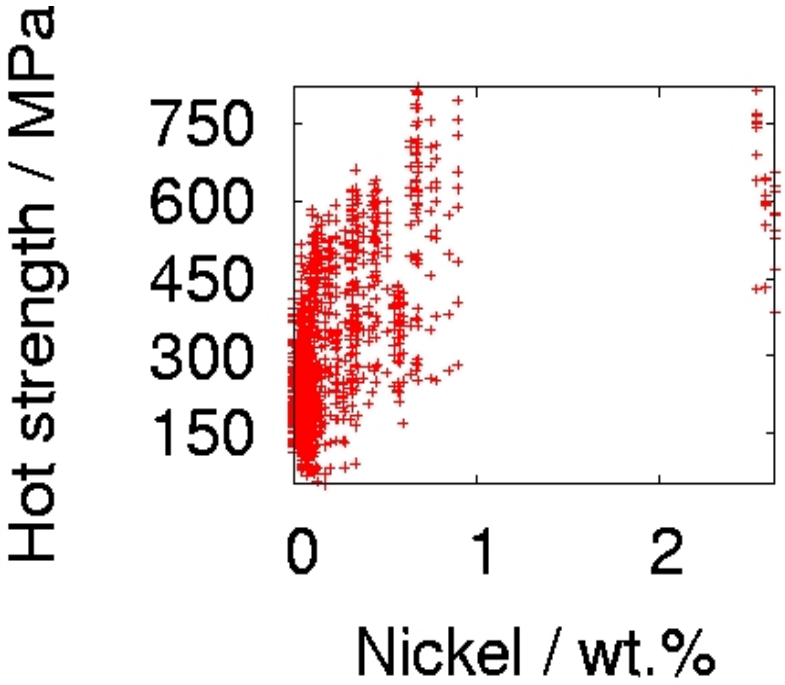
- THAT WILL ENSURE A RELIABLE PREDICTION

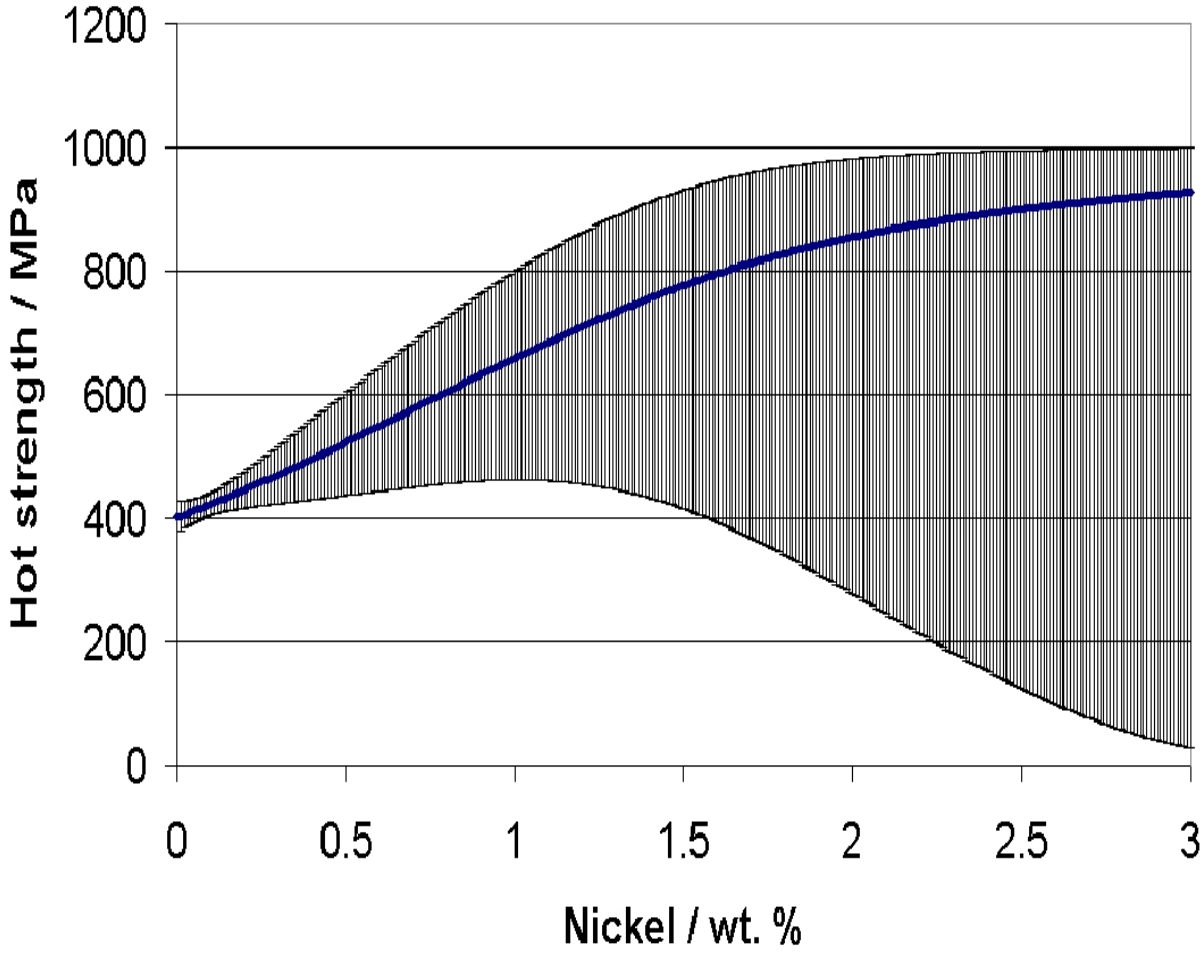
DATA-SPREAD AND PREDICTIONS OF HOT-STRENGTH AS A FUNCTION OF COPPER CONTENT



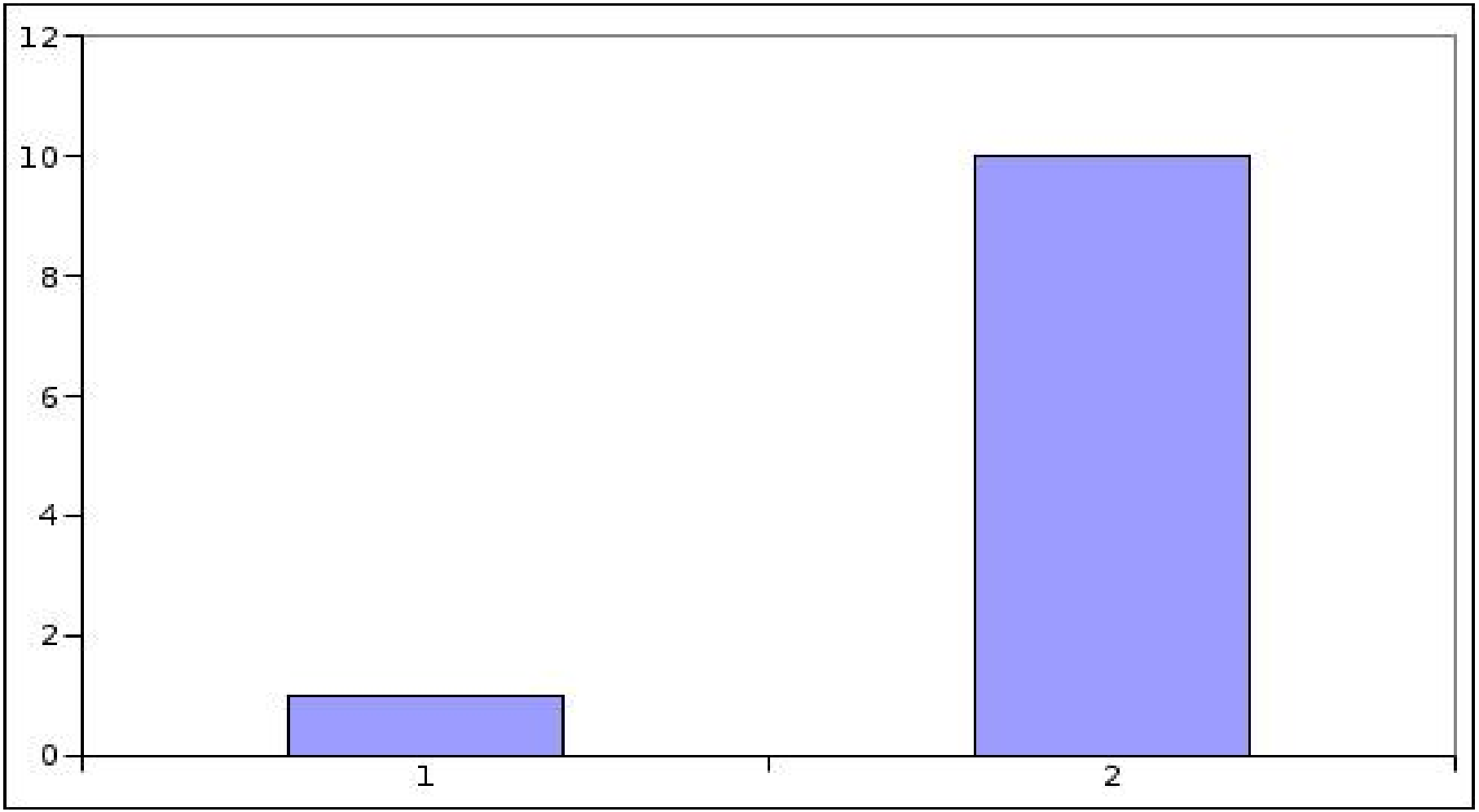


DATA-SPREAD AND PREDICTIONS OF HOT-STRENGTH AS A FUNCTION OF NICKEL CONTENT





CHECKING THE MINIMUM AND MAXIMUM



ALL THE INPUTS AND OUTPUTS

- MUST BE CHECKED BEFORE MODELLING
- TO ENSURE THAT THE LIMIT OF THE INPUTS AND OUTPUT IS MEANINGFUL

DO NOT TRUST THE EXPERIMENTS

- ALWAYS MAKE AN ASSESSMENT OF THE QUALITY OF THE DATA

- QUESTION THE ACCURACY AND RELIABILITY OF THE EXPERIMENTS USED TO GENERATE THE DATA

LOOK FOR OUTLIERS

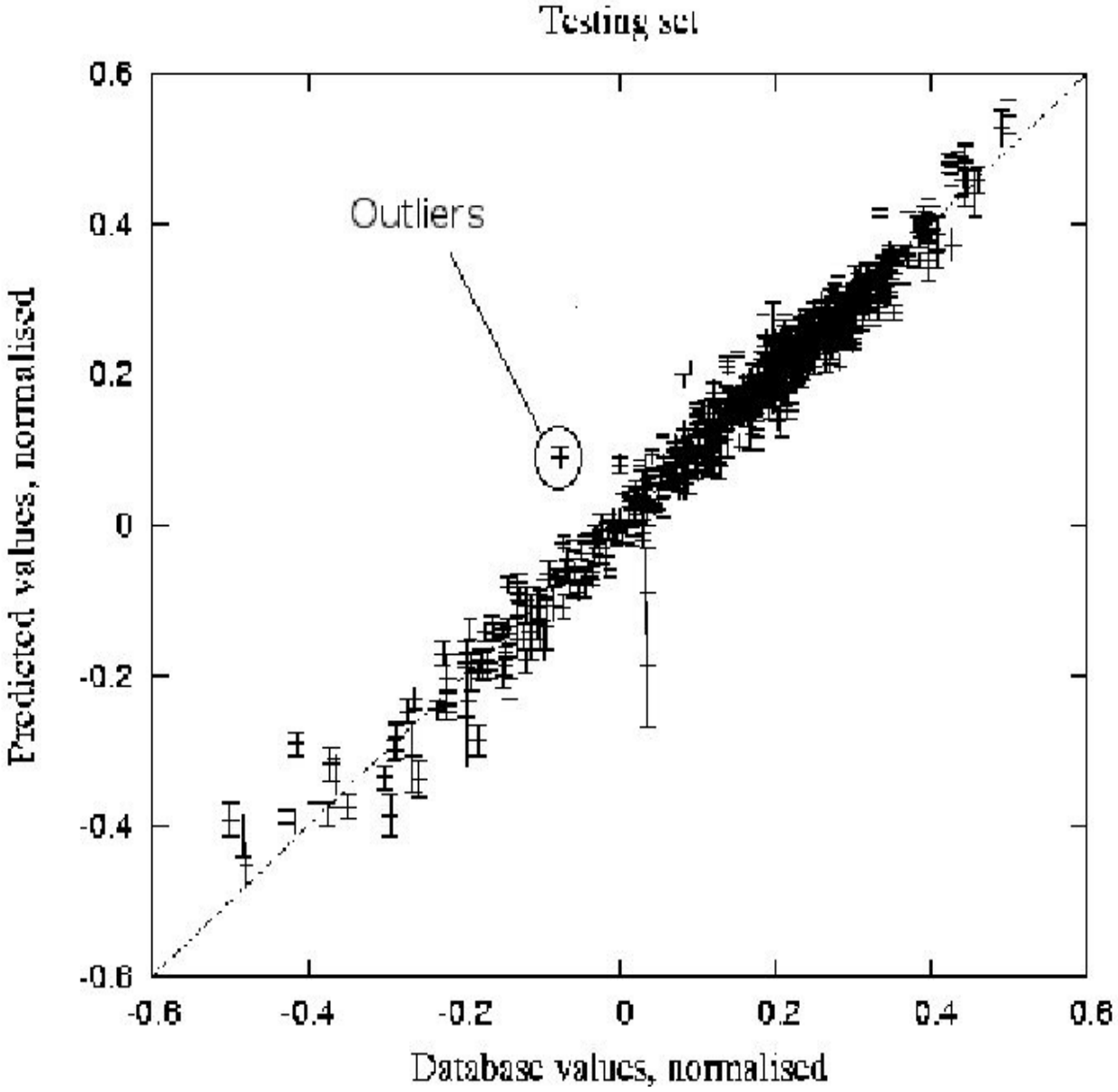
OUTLIERS

ARE POINTS

WHICH ARE

VERY

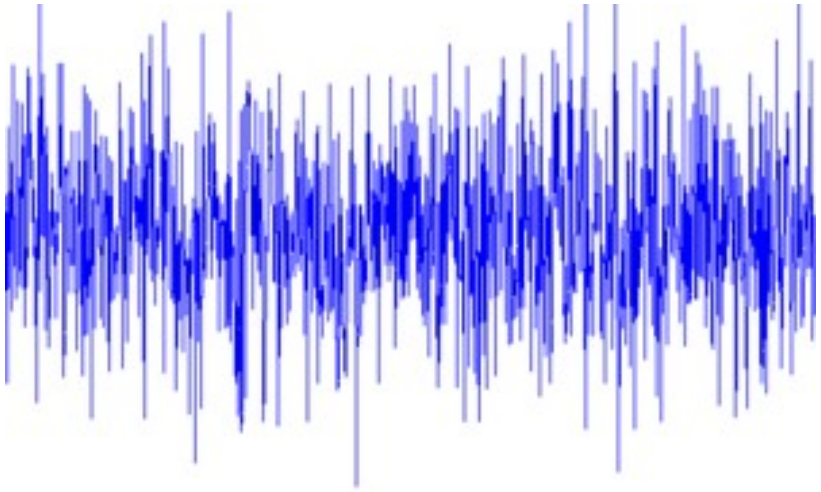
UNLIKELY



STUDY OUTLIERS

- FIRST THEY MUST BE IDENTIFIED
- TRACKED BACK TO THE ORIGIN TO LOOK FOR MISTAKES IN COLLECTING AND COMPILING THE DATABASE
- IF NO ERRORS ARE FOUND THEN SOME THINKING AND ASSUMING MUST BE DONNE TO ENSURE THE CORECTITUDE OF THE MODEL

FIXING THE NOISE LEVEL



BEFORE TRAINING

WE HAVE THE

POSSIBILITY

TO FIX THE

NOISE LEVEL

WE FIX THE LEVEL OF NOISE IF:

- WE KNOW WE HAVE A NOISY DATABASE
- AND IF WE TRAIN WITH A LOW LEVEL OF NOISE THE DATA WILL TRAIN UNTIL IT REACHES THE FIXED LEVEL (OVER-TRAIN)
- BY INCREASING THE LEVEL OF NOISE WE ENSURE A GOOD TRAINING

AND IF:

- WE KNOW WE HAVE A EXCELLENT DATABASE

- THEN WE LOWER THE NOISE LEVEL AND FORCE THE NEURAL NETWORK TO FIND A MORE COMPLEX SOLUTION THEN IT WOULD HAVE WITH A MEDIUM LEVEL

SO A GOOD DATABASE

- HAS INPUTS PARAMETERS SIGHIFICANT TO THE OUTPUT PARAMETER
- AS MANY SIGNIFICANT PARAMETERS AS POSSIBLE
- UNIFORMLY DISTRIBUTED DATA
- DOES NOT HAVE OUTLIERS
- HAS A OPTIMUM LEVEL OF NOISE

INPUT PARAMETERS

- WE CAN INSERT THE INPUT PARAMETERS AS RAW

FOR EXAMPLE

- THE CHEMICAL COMPOSITION OF A STEEL
- HEAT TREATMENT TEMPERATURE AND TIME
- GRAIN SIZE
- HARDNESS
- .....

OR

- WE CAN INSERT THE INPUTS IN A FUNCTIONAL FORM

FOR EXAMPLE

- THERMAL ACTIVATION
- LOGARITHM OF TIME
- A COMBINATION OF INPUT PARAMETERS
- NORMALISATION

THERMAL ACTIVATION

THE FORMULA DESCRIBES THE DEPENDENCE OF THE OUTPUT UPON AN ACTIVATION ENERGY

$$exp(-\frac{Q}{RT})$$

LOGARITHM OF TIME

IT IS WISE TO INCORPORATE THE LOGARITHM OF TIME IN THE DATABASE WHEN THE EXTENT OF REACTION IS EXPECTED TO VARY WITH

$$\ln\{\text{time}\}$$

A COMBINATION OF INPUT PARAMETERS

IF A COMBINATION OF ONE OR MORE PARAMETERS HAS A PARTICULAR SIGNIFICANCE THAN THAT COMBINATION SHOULD BE INSERTED IN THE DATABASE

FOR EXAMPLE

- THE FOLLOWING PRODUCT CAN BE USED TO AS AN INPUT FROM KINETIC THEORY, A FUNCTION OF TIME AND TEMPERATURE

$$time * exp(-\frac{Q}{RT})$$

NORMALISATION OF INPUTS

- USUALLY IT IS GOOD TO NORMALISE THE INPUTS IN ORDER TO MAKE COMPARISON BETWEEN THEM EASIER

- FROM THE COMPARISON WE CAN SEE WHICH INPUT PARAMETER HAS A BIGGER INFLUENCE ON THE MODEL

FOR EXAMPLE

- IF WE WANT TO COMPARE THE EFFECT OF CARBON WHICH VARIES BETWEEN 0.1-1 wt% AND CHROMIUM WHICH CAN VARY BETWEEN 0-13 wt%.

- IT IS EASY TO SEE THEIR EFFECT ON THE MODEL WHEN THEY ARE NORMALISED

BUT

- IT IS SENSITIVE THAT ALONG THE RAW DATA THE DATA THAT WAS TRANSFORMED AS A FUNCTIONAL FORM TO BE KEPT IN THE DATABASE

- IF WE DO NOT KEEP ALL THE DATA WE RUN THE RISK OF BIASING THE MODEL

THE OUTPUTS

- THE OUTPUTS CAN ALSO BE INSERTED IN A RAW FORM IN THE MODEL

OR

- AS A MATHEMATICAL FUNCTION FOR THE MODELS THAT HAVE THE TENDENCY TO PREDICT UNPHYSICAL VALUES

FOR EXAMPLE

WE HAVE THE STRENGTH AS AN OUTPUT AND AS A MATHEMATICAL FUNCTION THE STRENGTH CAN GO BELOW ZERO, THAT IS PHYSICALLY IMPOSSIBLE SO WE CAN GUIDE TO OUTPUT NOT TO REACH ZERO

$$f_{strength} = ln(-ln(1-\frac{X_{max}-X}{X_{max}-X_{min}}))$$

ACKNOWLEDGEMENTS

- MANY THANKS TO PROFESSOR H. K. D. H. BHADESHIA
- DEPARTMENT OF MATERIAL SCIENCE AND METALLURGY
- UNIVERSITY OF CAMBRIDGE
- EUROPEAN COMMISSION UNDER THE MARIE CURIE EARLY STAGE RESEARCH TRAINING PROGRAMME