Walla Walla University
SEVENTH-DAY ADVENTIST HIGHER EDUCATION

North American Simulation and Gaming Association
NASAGA

www. Mihail.Motzev.com

ABSEL
Association for Business Simulation and Experiential Learning

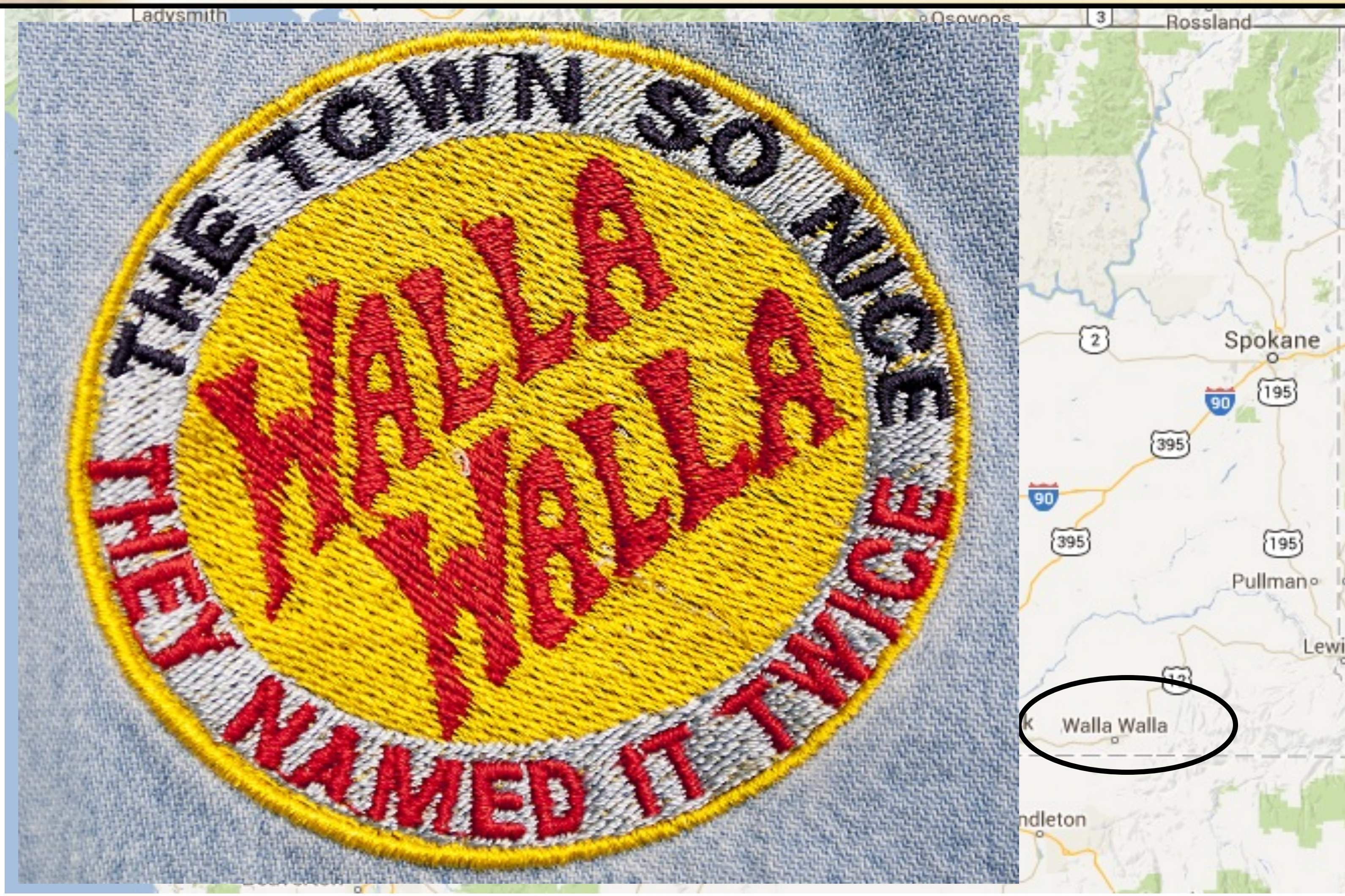
ISAGA

Summer school on research methods ‘2020:

“STATISTICAL LEARNING NETWORKS”

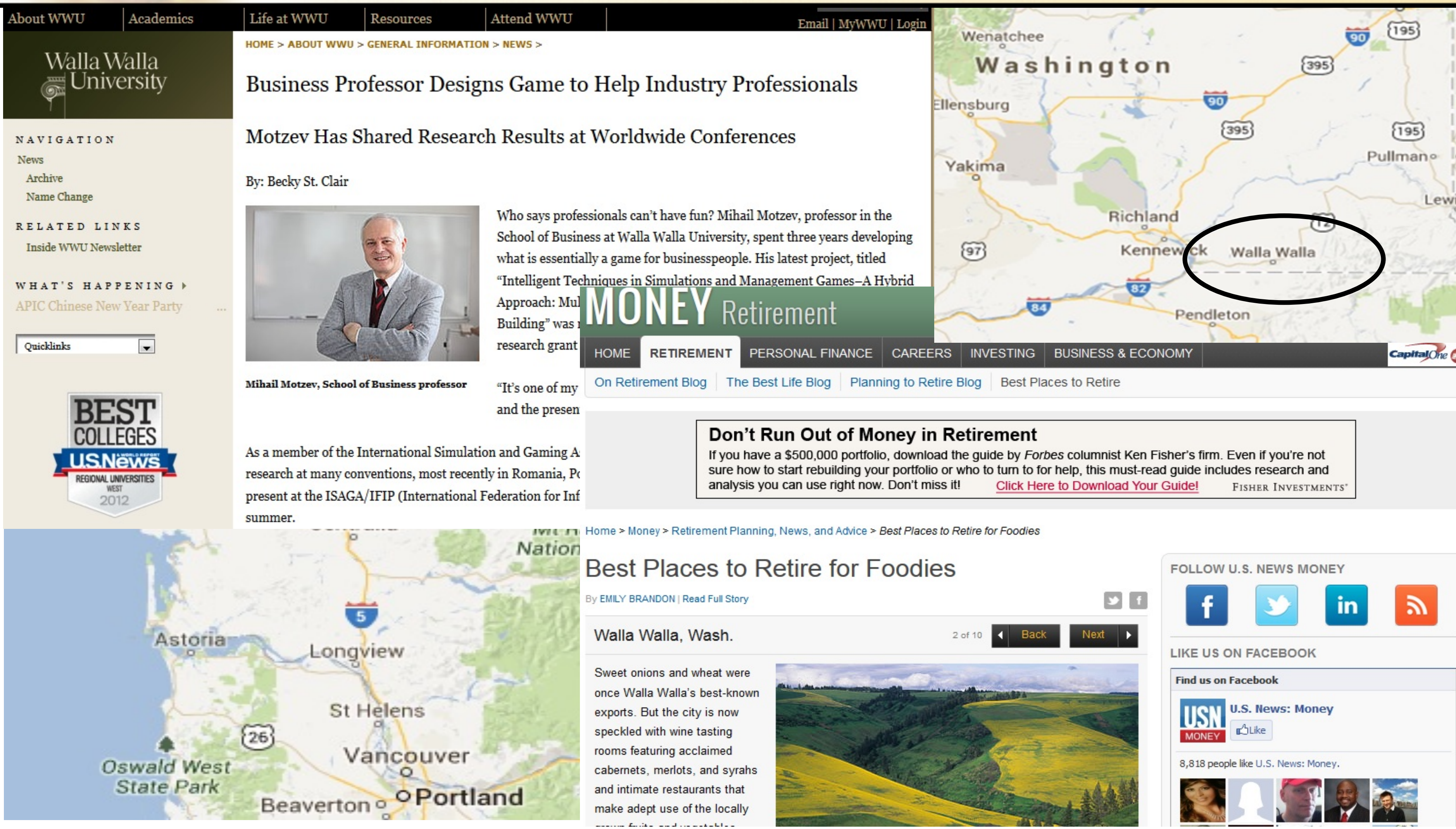
Mihail Motzev
Ph.D,M.Sc,P.D.D
(M.Motzev@wallawalla.edu)

Greetings from Walla Walla



Walla Walla University
SEVENTH-DAY ADVENTIST HIGHER EDUCATION

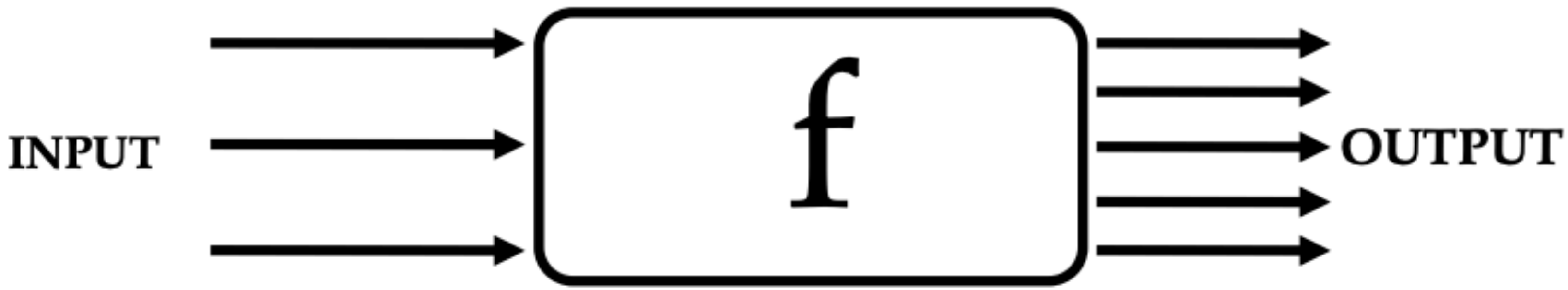
A Town Offering Something for Everyone



Learning Models & Approaches

- ***Supervised learning*** **- is the machine learning task of learning a function that maps an input to an output based on example input-output pairs.**
- ***Unsupervised learning*** **- looks for previously undetected patterns in a data set with no pre-existing labels and with a minimum of human supervision, also known as self-organization.**
- ***Semi-supervised learning*** **- an approach to machine learning that combines a small amount of labeled data with a large amount of unlabeled data during training.**

Statistical Learning Theory: supervised learning



Given a set of l examples (data)

$$\{(x_1, y_1), (x_2, y_2), \ldots, (x_\ell, y_\ell)\}$$

Question: find function f such that

$$f(x) = \hat{y}$$

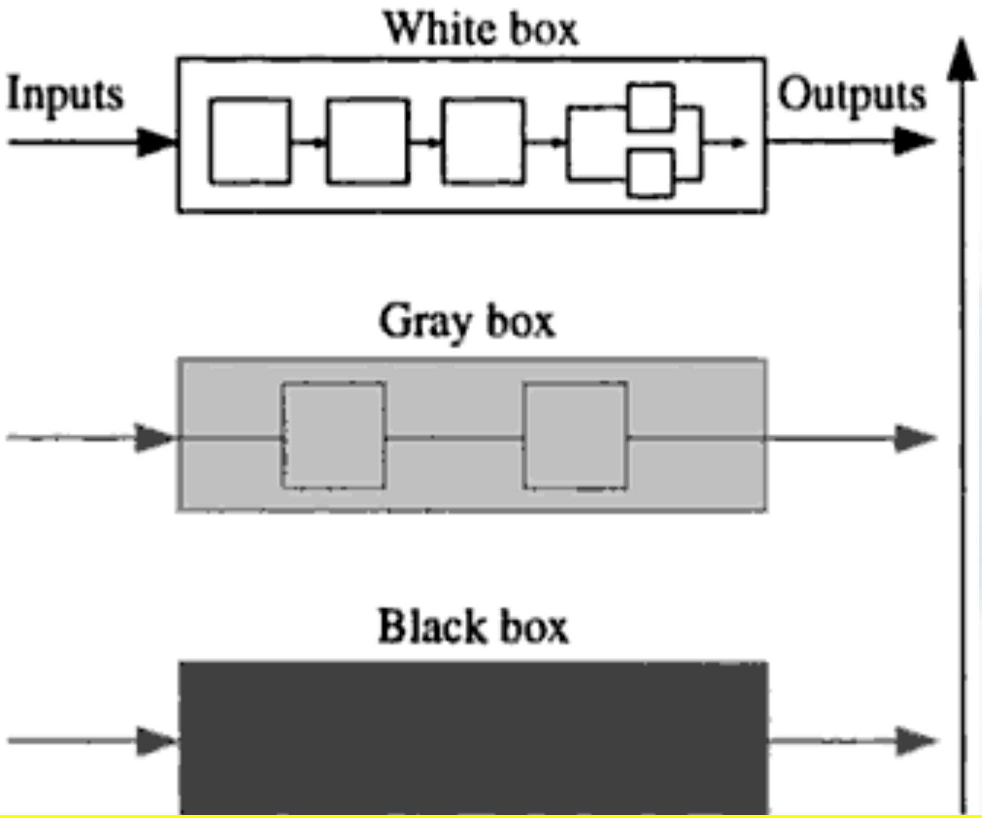
is a **good predictor** of y for a **future** input x (fitting the data is **not** enough!)

A framework for machine learning drawing from the fields of statistics and functional analysis.

Fundamentals:

- ***Network – a function (model) represented by the composition of many basic functions (models).***
- ***Basic function – element, unit, building block, network node, artificial neuron, partial model.***
- ***A Learning Network estimates its function from representative observations of the relevant variables.***
- ***From a data mining perspective, ANNs are just another way of fitting a model to observed historical data in order to be able to make classifications or predictions.***

Model Identification



Increasing internal knowledge

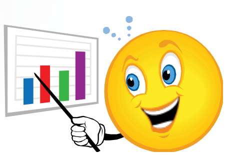
Data Mining

Computational process of discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems

Data Mining and Predictive analytics

- ***Data mining*** is the process of exploration and analysis (by automatic or semi- automatic means) of large quantities of data in order to discover meaningful patterns and rules.

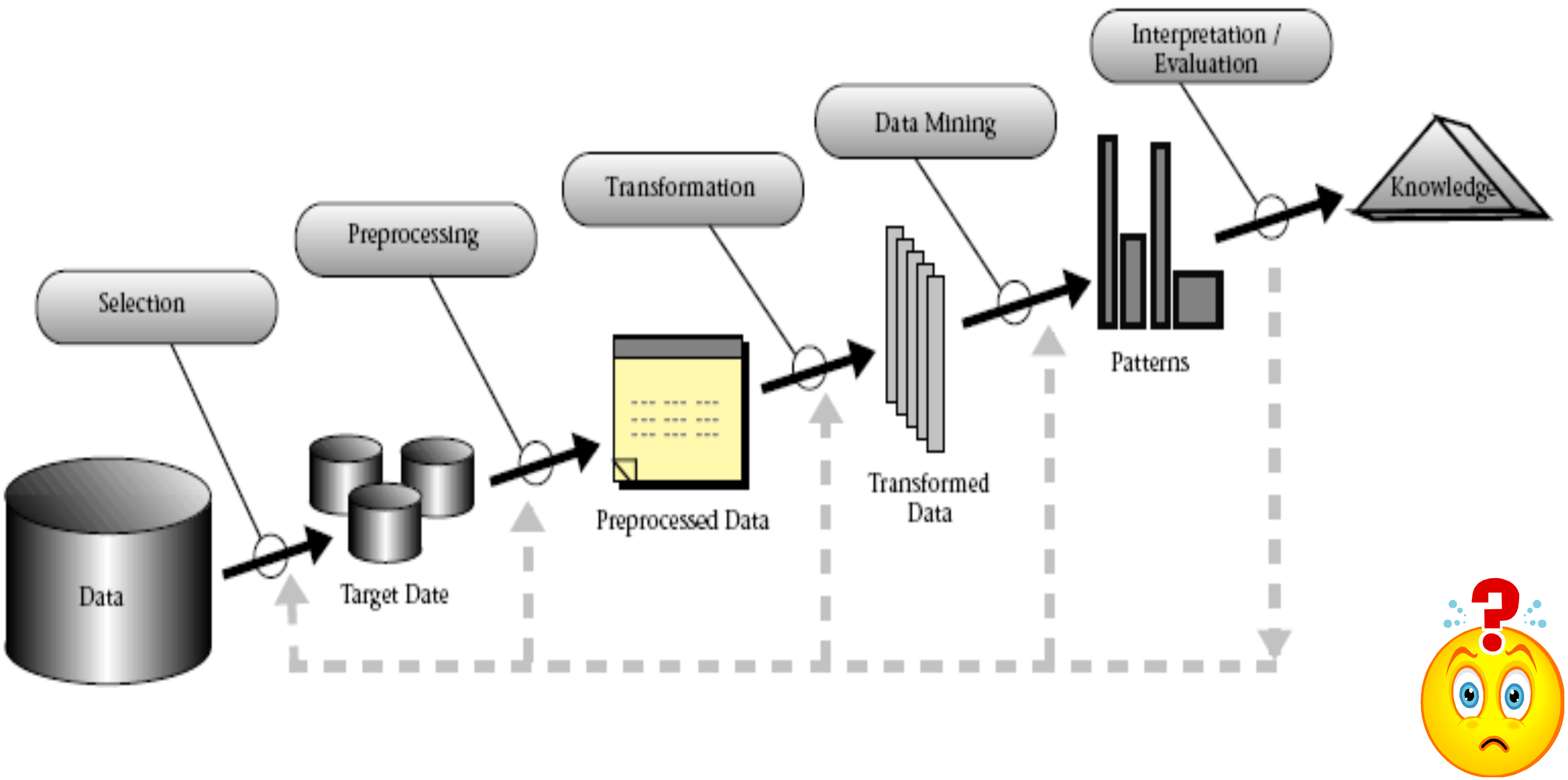
- ***Predictive analytics*** encompasses a variety of techniques from *statistics*, *machine learning* and *data mining* that analyze current and historical facts to make predictions about future or otherwise unknown events - technically, predictive analytics is an area of data mining that deals with extracting information from data and using it to predict trends and behavior patterns.



Knowledge Discovery in Databases -

Identification of underlying patterns, categories, and behaviors in large data sets, using techniques such as neural networks and data mining

An overview of KDD process

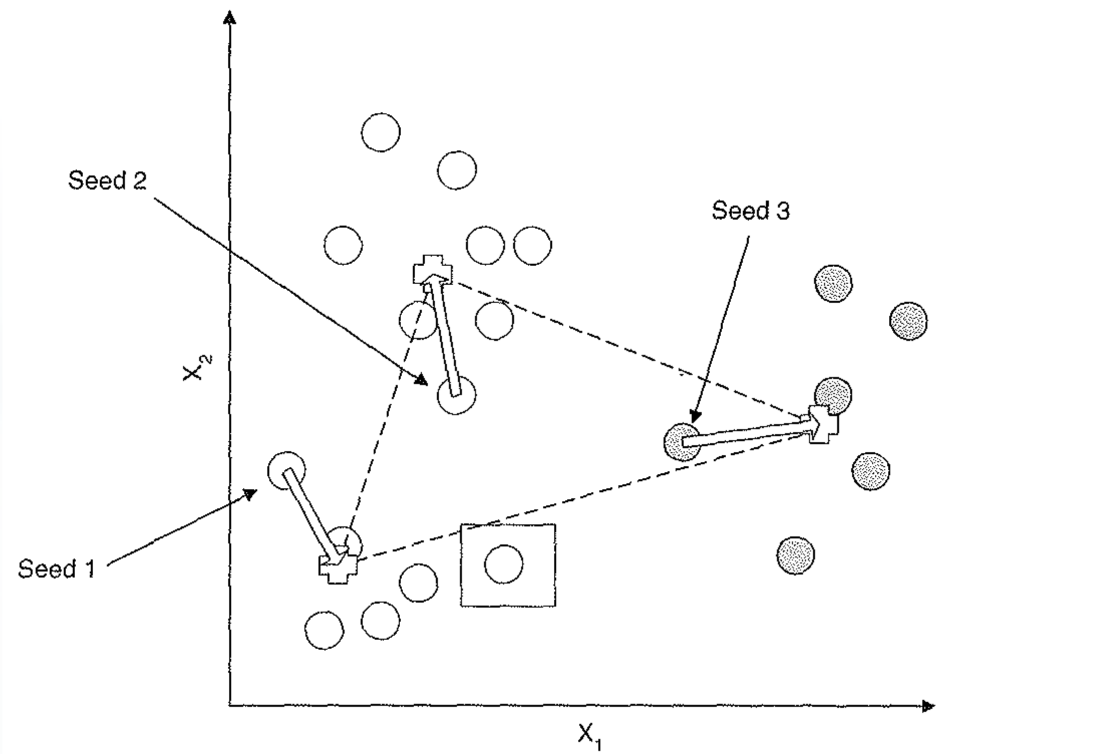


Data Mining Techniques

- ***Automatic Cluster Detection -*** **use cluster detection when we suspect that there are natural groupings that may represent groups of customers or products that have a lot in common with each other.**
- ***Decision Trees (Classification & Regression) -*** **a good choice when the data mining task is classification of records or prediction of outcomes. We should use decision trees when the goal is to assign each record to one of a few broad categories.**
- ***Artificial Neural Networks (the most widely known and the least understood of the major data mining techniques) -*** **a good choice for most classification and prediction tasks when the results of the model are more important than understanding how the model works. *ANNs* represent complex mathematical equations, with lots of summations, exponential functions, and many parameters.**

Automatic Cluster Detection

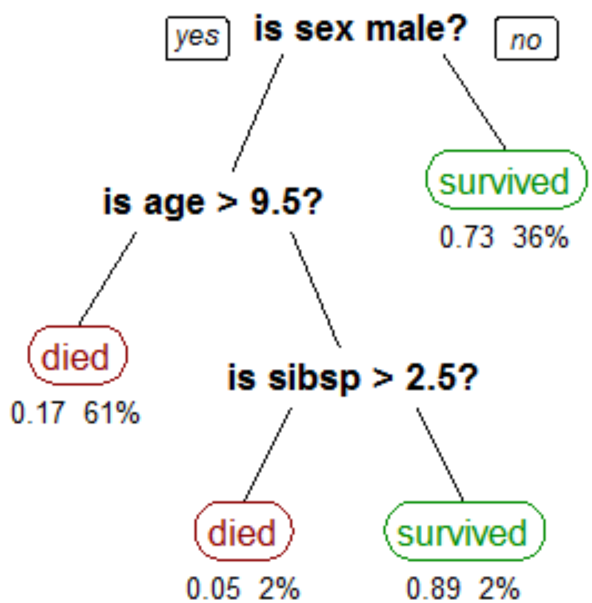
Data Mining Techniques



Grouping a set of objects in such a way that objects in the same group (cluster) are more similar (in some sense or another) to each other than to those in other groups (clusters)

Decision Trees

Data Mining Techniques

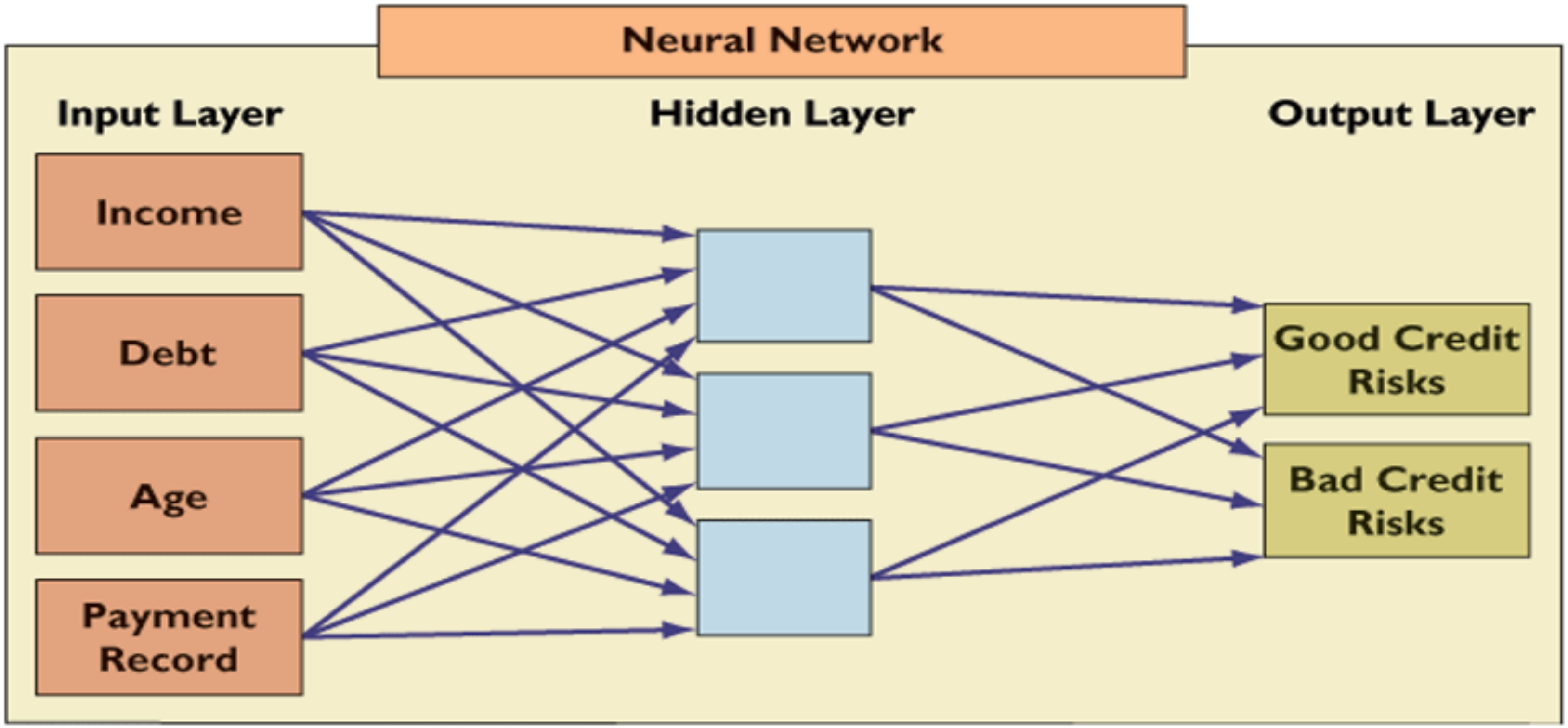


A tree showing survival of passengers on the Titanic ("sibsp" is the number of spouses or siblings aboard). The figures under the leaves show the probability of survival and the percentage of observations in the leaf

Artificial Neural Networks

Data Mining Techniques

ANN – artificial systems which emulate the processing patterns of the biological brain to discover patterns and relationships in massive amounts of data *("Perceptron" - Ph. Rozenblat)*

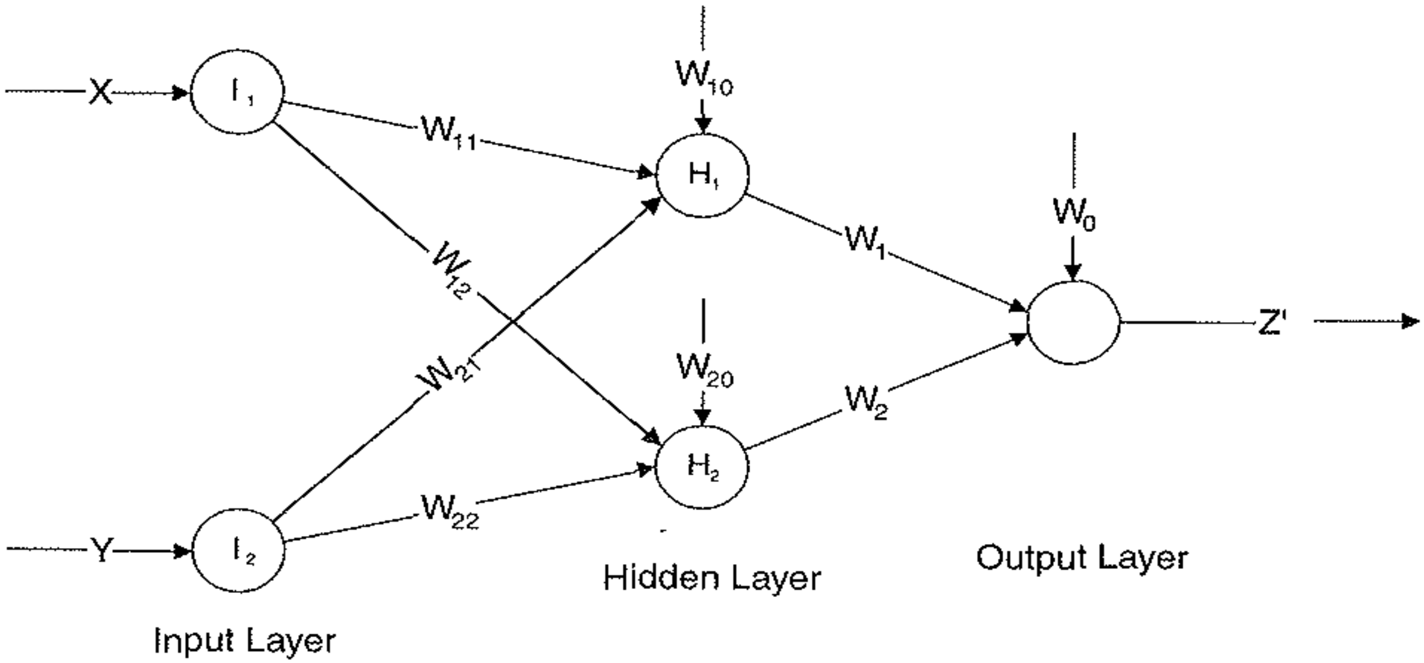


Source: Herb Edelstein, "Technology How-To: Mining Data Warehouses," *InformationWeek*, January 8, 1996. Copyright © 1996 CMP Media, Inc., 600 Community Drive, Manhasset, NY 11030. Reprinted with permission.

Walla Walla University
SEVENTH-DAY ADVENTIST HIGHER EDUCATION

Model Building

Data Mining Techniques - ANNs

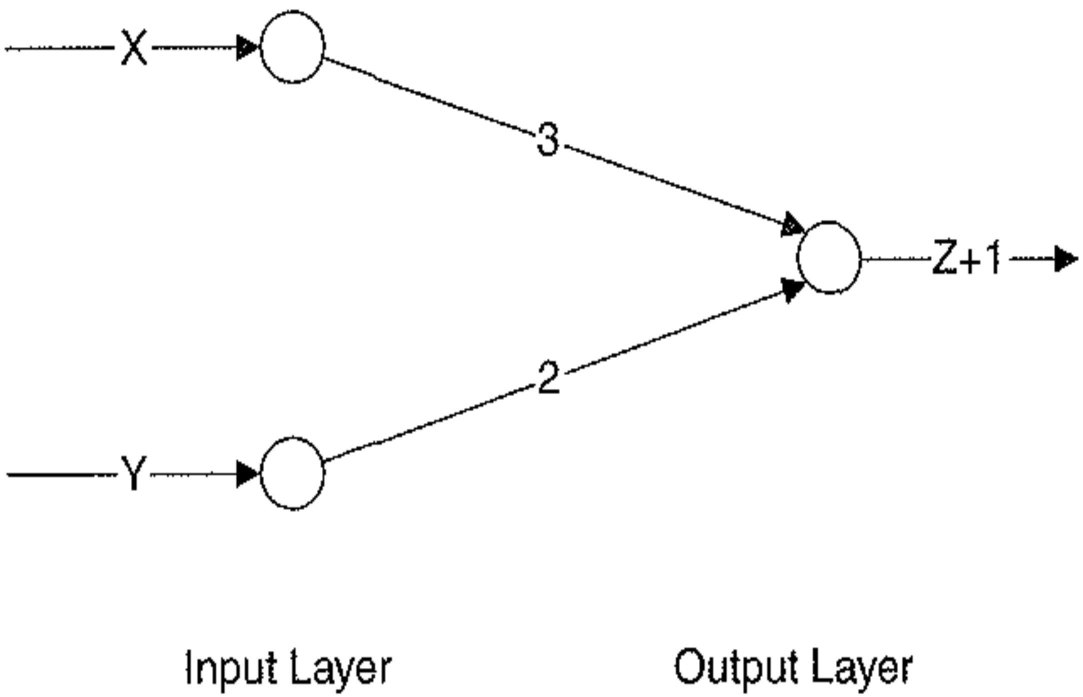


A neural network with a hidden layer.

"The most widely known and the least understood of the major data mining techniques."

How a Neural Network Works

Data Mining Techniques



Neural network representation of $z=3x+2y-1$.

Data Mining and Predictive Modelling

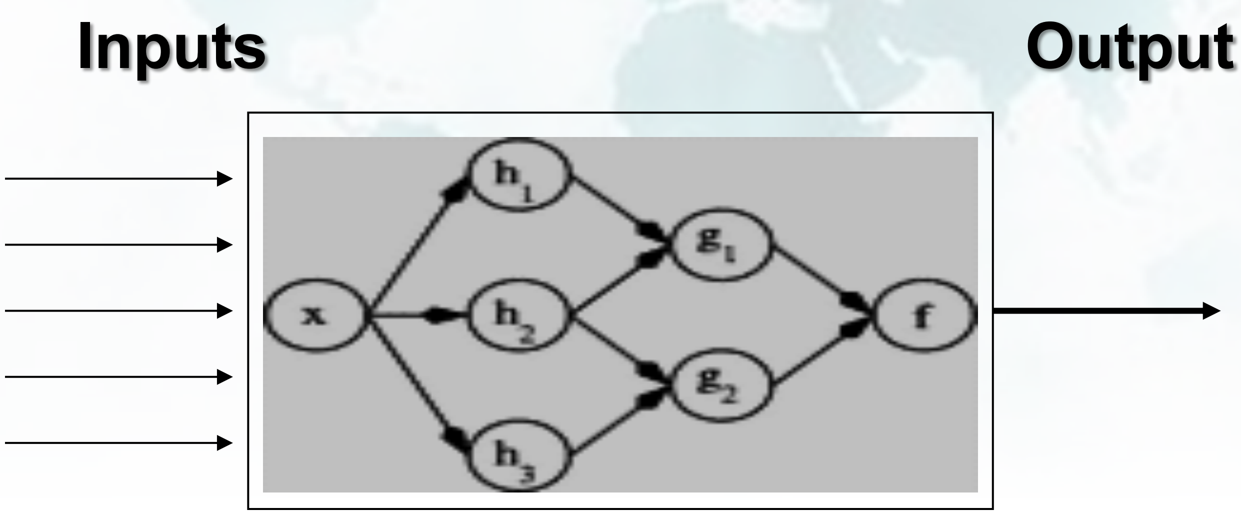
Directed Data Mining

- *A top-down approach* – often takes the form of predictive modeling where we know exactly what we want to predict. In this case the model is considered as a *black box*, i.e. it is not important what the model is doing, we just want the most accurate result possible.

Data Mining and Predictive Modelling

Undirected Data Mining

- ***A bottom-up approach*** **that finds patterns in the data which provide insights. This form of data mining is represented with *semitransparent boxes* and unlike directed *DM*, here users want to know what is going on, how the model is coming up with an answer.**



	ANNs	Statistical Learning Networks
Data analysis	universal approximator	structure identificator
Analytical model	indirect by approximation	direct
Architecture	unbounded network structure; experimental selection of adequate architecture demands time and experience	bounded network structure [1]; adaptively synthesised structure
A-priori-Information	without transformation in the world of ANNs not usable	can be used directly to select the reference functions and criteria
Self-organisation	deductive, given number of layers and number of nodes (subjective choice)	inductive, number of layers and of nodes estimated by minimum of external criterion (objective choice)
Parameter estimation	in a recursive way; demands long samples	estimation on training set by means of maximum likelihood techniques, selection on testing set (extremely short)
Feature	result depends on initial solution, time-consuming technique, necessary knowledge about the theory of neural networks	existence of a model of optimal complexity, not time-consuming technique, necessary knowledge about the task (criteria) and class of system (linear, non-linear)

General Prediction Model

$$y = a_0 + \sum_{i=1}^{M} a_i x_i + \sum_{i=1}^{M}\sum_{j=1}^{M} a_{ij} x_i x_j + \sum_{i=1}^{M}\sum_{j=1}^{M}\sum_{k=1}^{M} a_{ijk} x_i x_j x_k$$

where $X(x_1, x_2, \ldots, x_M)$ - input variables vector;

$A(a_1, a_2, \ldots, a_M)$- vector of coefficients or weights.

$$Y = F(X, e)$$

where ***F*** can be any mathematical function describing the variable ***Y (the output)*** as a function of ***input variables X*** and the stochastic component ***e (model error).***

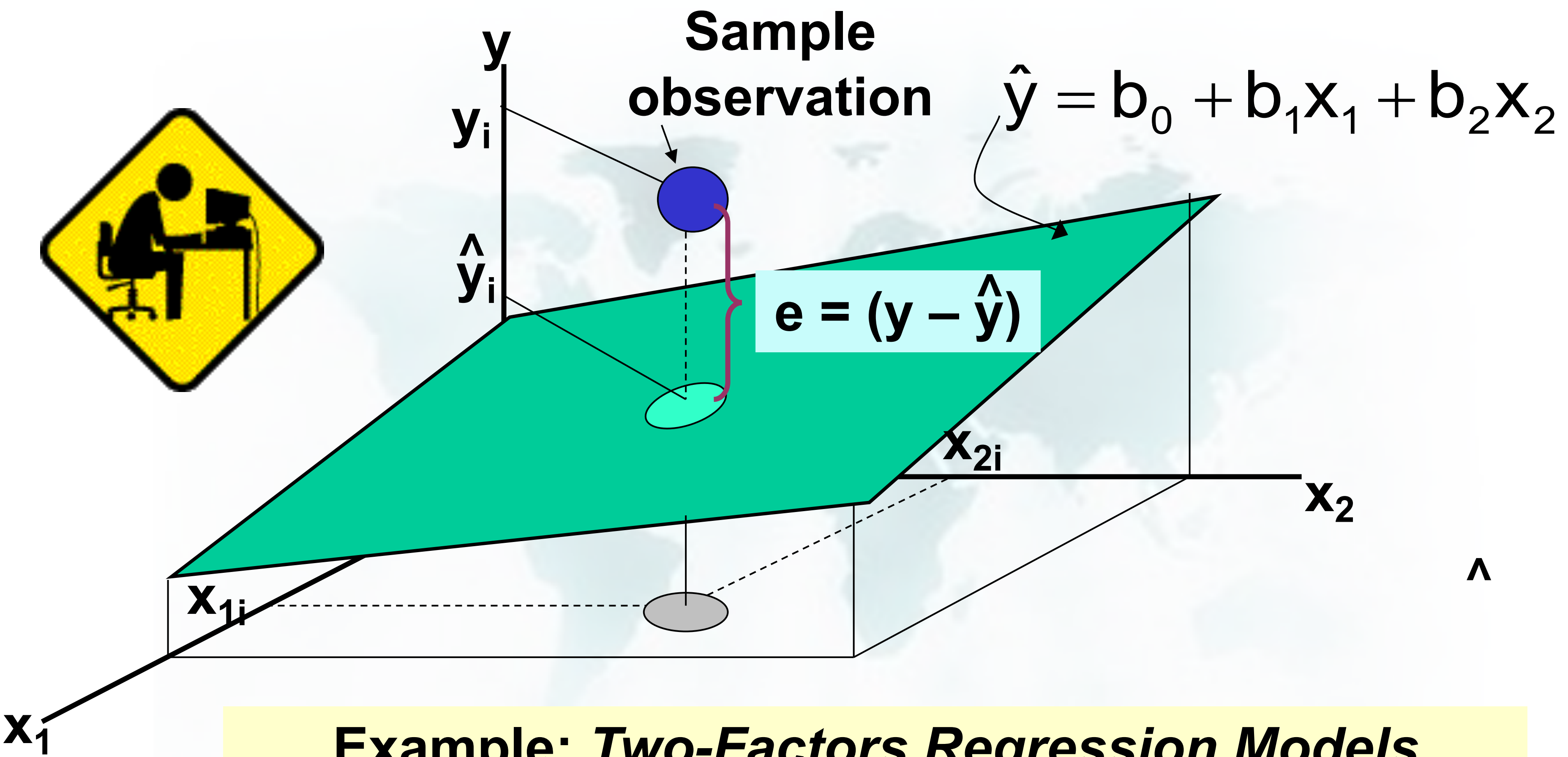
Model Building

Problems

- *Model specification;*
- *Overfitting;*
- *Autocorrelation;*
- *Multicollinearity*
- *ANNs:*
  - *- number of layers;*
  - *- how many input nodes;*
  - *- best activation function;*
  - *- ANN training;*
  - *- lack of transparency (interpretation), etc.*

Model Building

Regression Analysis



Example: ***Two-Factors Regression Models***

Model Building

Regression Models – Problems:

Alan Greenspan (The Map and the Territory: Risk, Human Nature, and the Future of Forecasting):

APPENDICES

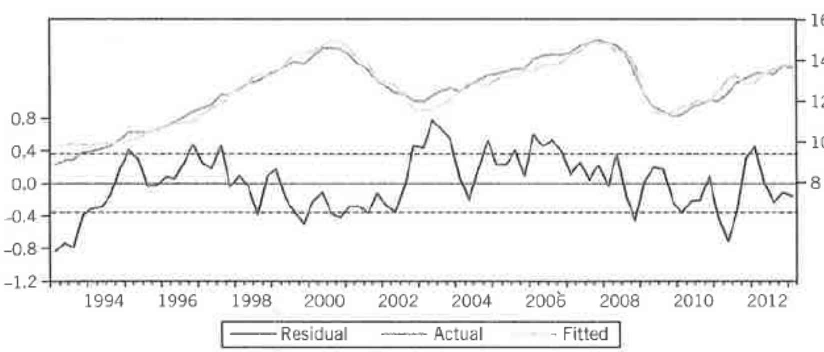
Exhibit 4.7

Dependent Variable (Time Period: Q1 1993–Q1 2013, 81 obs.)
Real Pvt Nonres Fixed Invst (SAAR, Bil.Chn.2005$) / Pvt Nonres Fixed Assets (2005 = 100)

Independent Variable(s)	Coefficient	t-Statistic*
S&P 500 (1941-43=10) / Pvt Nonres Fixed Invst Price (SA, 2005 = 100) (1 quarter ago)	0.473	19.044
Nonfarm Operating Rate (SA, % of capacity) (3 quarters ago)	0.165	6.118
Structures' share of nominal Pvt Nonres Fixed Invst	6.332	4.517

Adjusted R-sq 0.946 — Durbin-Watson 0.585

*t-statistic calculated using Newey-West HAC standard errors and covariance.



Source: U.S. Department of Commerce; Standard and Poor's; Federal Reserve Board; author's calculations.

Exhibit 3.3

Dependent Variable (Time Period: Jan. 1991–Dec. 2005, 180 obs.
m/m %Δ in: CoreLogic Home Price Index (Seasonally adjusted

Independent Variable(s)
Freddie Mac 30yr Fixed-Rate Mortgage Rate, % p.a. (3 mo

Adjusted R-sq 0.604 — Durbin-Watson 0.159

*t-statistic calculated using Newey-West HAC standard errors and covarian

Exhibit 4.6

Dependent Variable (Time Period: Q1 1970–Q4 2012, 172 obs.)
ln [Real GDP / Real GDP (4 quarters ago)]

Independent Variable(s)	Coefficient	t-Statistic*
ln [**Corp & Home Equity, Period Avg (1 quarter ago) / **Corp & Home Equity, Period Avg (5 quarters ago)]	0.127	9.691

Adjusted R-sq 0.419 — Durbin-Watson 0.364

*t-statistic calculated using Newey-West HAC standard errors and covariance.
**Domestic holdings of domestic corporate equities and foreign corporate equities, at market value.

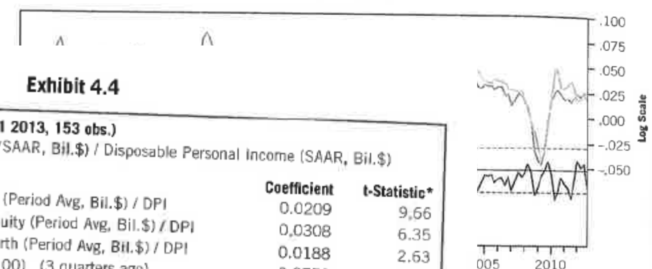


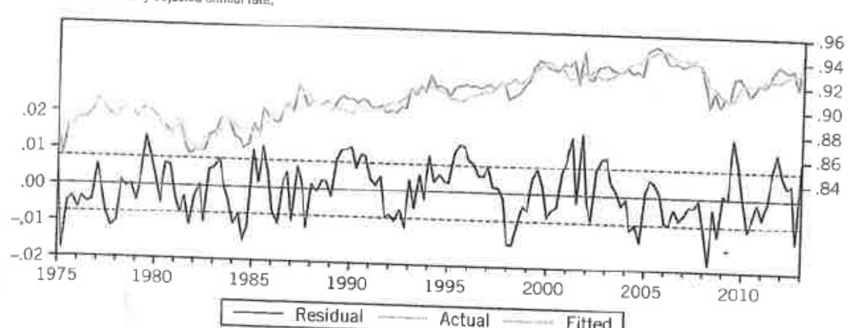
Exhibit 4.4

Dependent Variable (Time Period: Q1 1975–Q1 2013, 153 obs.)
Personal Consumption Expenditures (***SAAR, Bil.$) / Disposable Personal Income (SAAR, Bil.$)

Independent Variable(s)	Coefficient	t-Statistic*
Household (incl. NPOs) Stock Net Worth (Period Avg, Bil.$) / DPI	0.0209	9.66
Household (incl. NPOs) Homeowners' Equity (Period Avg, Bil.$) / DPI	0.0308	6.35
Household (incl. NPOs) All Other Net Worth (Period Avg, Bil.$) / DPI	0.0188	2.63
6-Month Certificates of Deposit (% p.a./100) (3 quarters ago)	−0.3752	−9.56
[**Adjusted PI / DPI] (2 quarters ago)	0.2666	2.30

Adjusted R-sq 0.903 — Durbin-Watson 1.089

*t-statistic calculated using Newey-West HAC standard errors and covariance.
**Adjusted PI = (0.9*Wages and Salary Disbursements) + (1.0*Personal Current Transfer Receipts) + (0.5*All Other Personal Income).
***Seasonally adjusted annual rate.



Source: Federal Reserve Board; U.S. Department of Commerce.

Model Building

Machine Learning - Interpretations

Simple numerical example

Consider the following data set :

y	a	b	c
9	1	8	1
9	2	7	2
9	3	6	3
9	4	5	4
9	5	4	5
9	6	3	6
9	7	2	7
6	99	1	5

Model:

Y = F (a,b,c)

Solutions:

y = 9.3- 0.033a- 0.033b

y = 0.00001 + b + c

y = 9 - 0.0319a + 0.0319c

Statistical Learning Networks of Active Neurons

GMDH

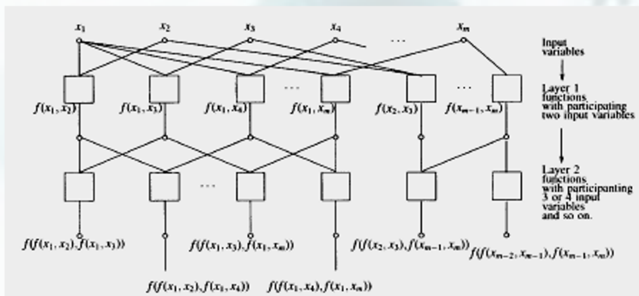
Gödel's incompleteness theorems

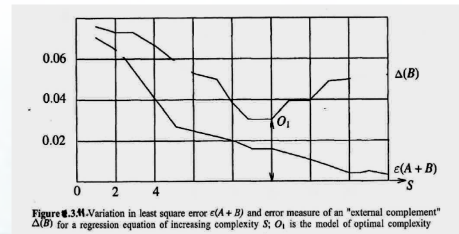
Alexey G. Ivakhnenko.

(1913-2007)

Two State Prizes of the USSR, Medal “For Labor”, Order of Friendship of Peoples …

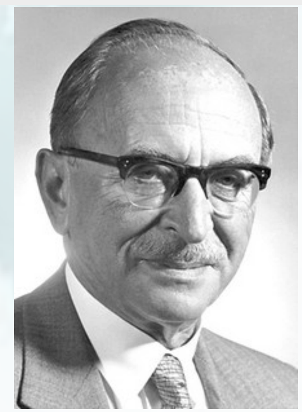
Genetic selection of pairwise features





Gabor’s principle of "freedom of decisions choice"

Knowledge extraction from experimental data, Self-Organization etc…



Dennis Gabor (1900-1978)

Numerous (>20) awards:

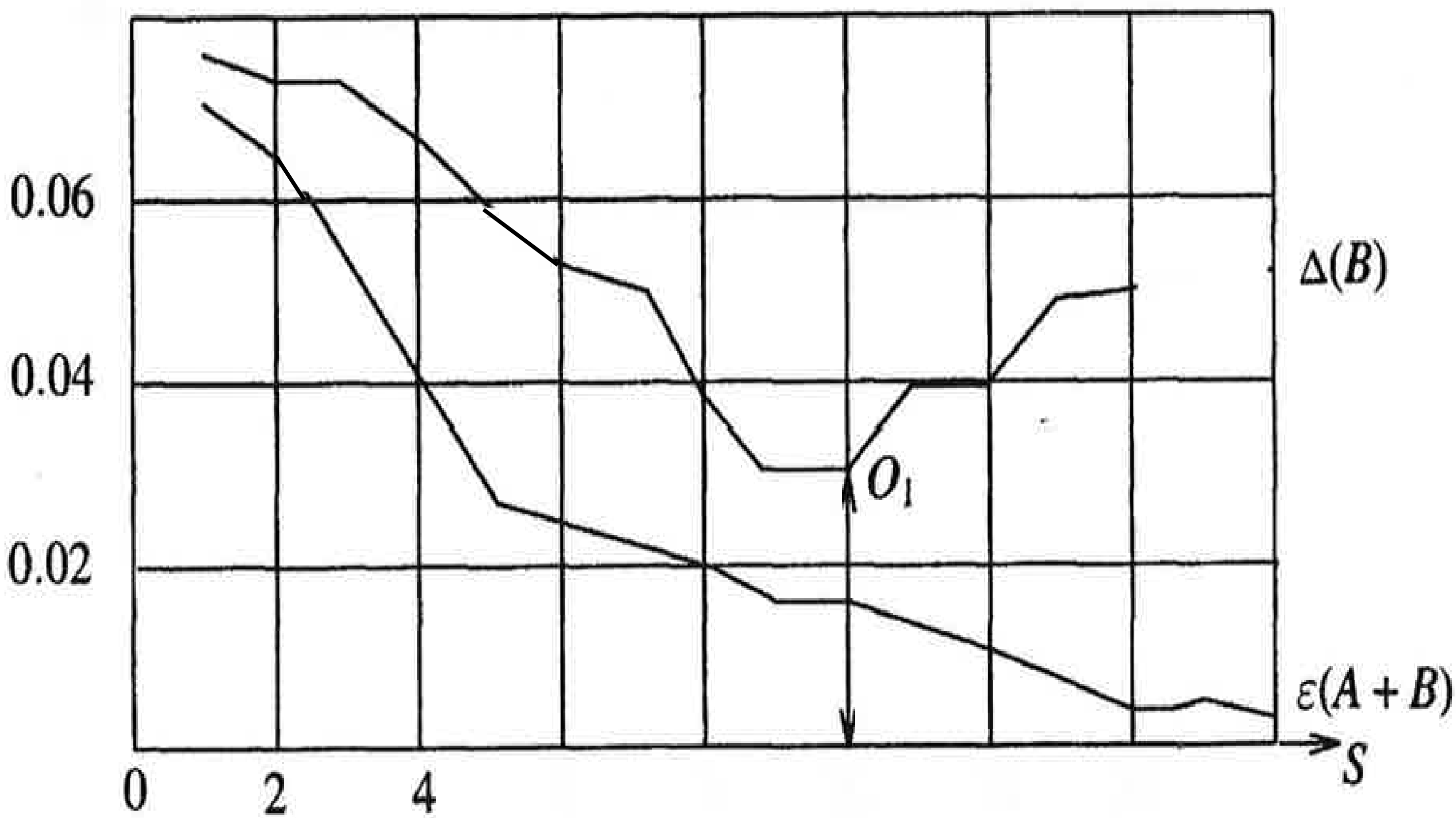
- Nobel Prize in Physics (1971)
- Honorary Doctorate, Delft University of Technology (1971)

Kurt Gödel (1906-1978)

Notable awards:

- Albert Einstein Award (1951)
- National Medal of Science (USA) in Mathematical, Statistical, and Computational Sciences (1974)

Overfitting – Internal vs External (Cross) Validation



Variation in least square error $\varepsilon(A+B)$ and error measure of an "external complement" $\Delta(B)$ for a regression equation of increasing complexity S; O_1 is the model of optimal complexity

Overfitting and Model Selection

- ***Cross Validation -*** also called ***rotation estimation*** or ***out-of-sample testing***, is a model validation technique for assessing how the results of a statistical analysis will generalize to an independent data set.
- Involves partitioning a sample of data into complementary subsets, performing the analysis on one subset (called the ***training set***), and validating the analysis on the other subset (called the ***validation set*** or ***testing set***).
- Two types of cross-validation can be distinguished: ***exhaustive*** and ***non-exhaustive cross-validation.***

Cross Validation

- ***Exhaustive cross-validation*** **- learn and test on all possible ways to divide the original sample into a training and a validation set.**

- *Leave-p-out cross-validation* *- involves using p observations as the validation set and the remaining observations as the training set. This is repeated on all ways to cut the original sample on a validation set of p observations and a training set.*

- ***Leave-one-out cross-validation*** - a particular case of ***leave-p-out cross-validation*** with p = 1.

Cross Validation

- ***Leave-one-out cross-validation:***

1. Select (it could be random) observation i for the testing set and use the remaining observations in the training set. Compute the error on the test observation.

2. Repeat the above step for i = 1, 2, … N-1, where N is the total number of observations.

3. Compute the forecast accuracy measures based on all errors obtained.

A total of 8 models will be trained and tested:

Cross Validation

- ***Non-exhaustive cross-validation* -** do not compute all ways of splitting the original sample. Those methods are approximations of ***leave-p-out cross-validation.***

- *k-fold cross-validation* - the sample is randomly partitioned into k equal sized subsamples. When k = n (the number of observations), *k-fold cross-validation* is equivalent to *leave-one-out cross-validation*.

- ***holdout method* -** randomly assign data points to two sets A and B (training set and test set).

- ***repeated random sub-sampling validation*** or ***Monte Carlo cross-validation*** creates multiple random splits of the dataset into training and validation data

Cross Validation

- ***Nested cross-validation*** **-** cross-validation is used simultaneously for selection of the best set of hyperparameters and for error estimation.

- *k*l-fold cross-validation* - contains an outer loop of k folds and an inner loop of l folds. One by one, a set is selected as (outer) test set and the k - 1 other sets are combined into the corresponding outer training set.

- ***k-fold cross-validation with validation and test set*** **-** k*l-fold cross-validation when l = k - 1. One by one, a set is selected as a test set. Then, one by one, one of the remaining sets is used as a validation set and the other k - 2 sets are used as training sets until all possible combinations have been evaluated.

Cross Validation with Time Series data

- ***Rolling forecasting origin*** **-** since it is not possible to get a reliable forecast based on a very small training set, the earliest observations ***n*** are not considered as testing sets.

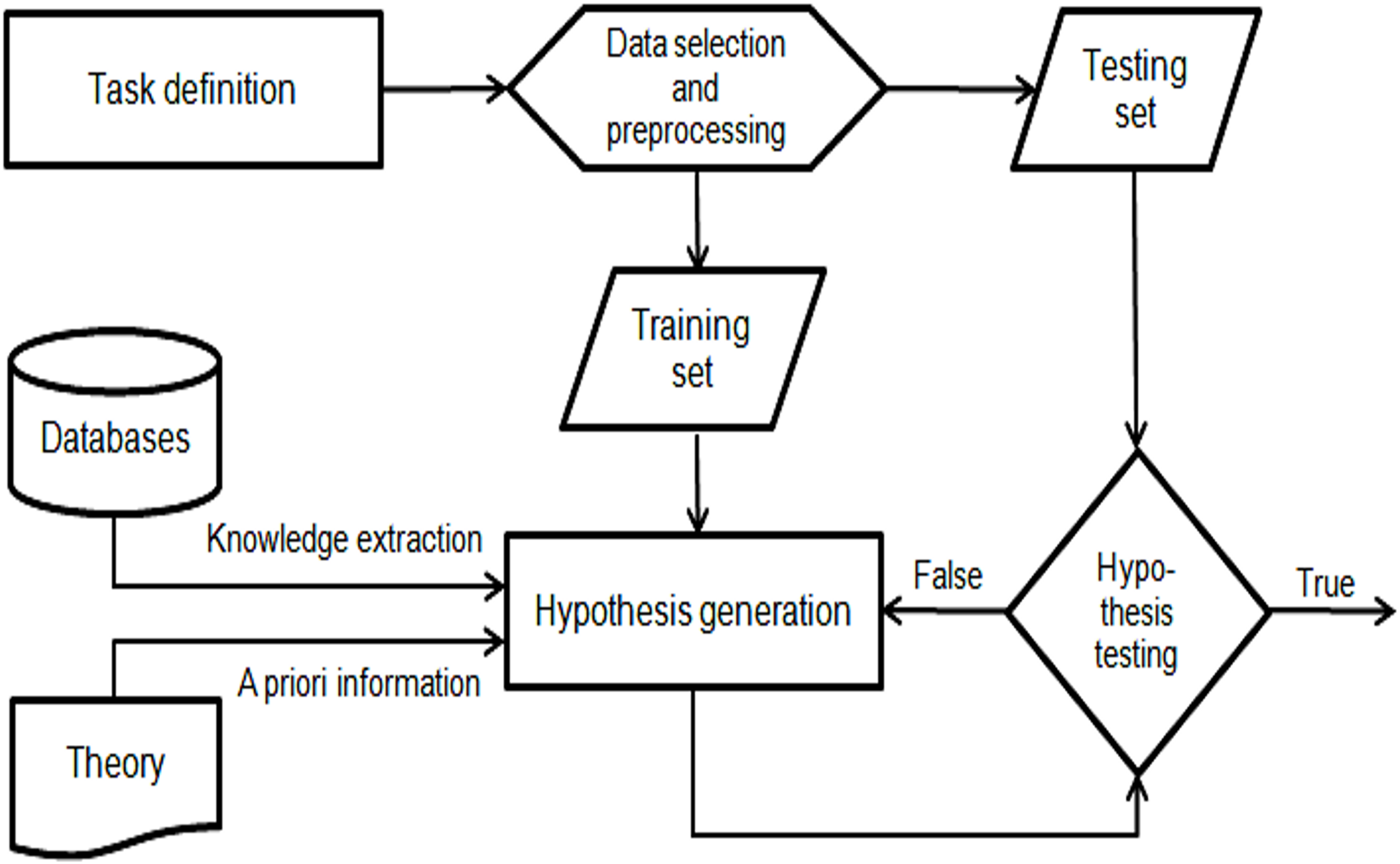
1. We select the observation at time (n+i) for the testing set and use the observations at times t= {1, 2, … (n+i-1)} to estimate the forecasting model. Then we compute the error on the forecast for the time (n+i).

2. The above step should be done for all i= {1, 2, … (T-n)}, where T is the total number of observations and the forecast error should be measured on each (n+i) period accordingly.

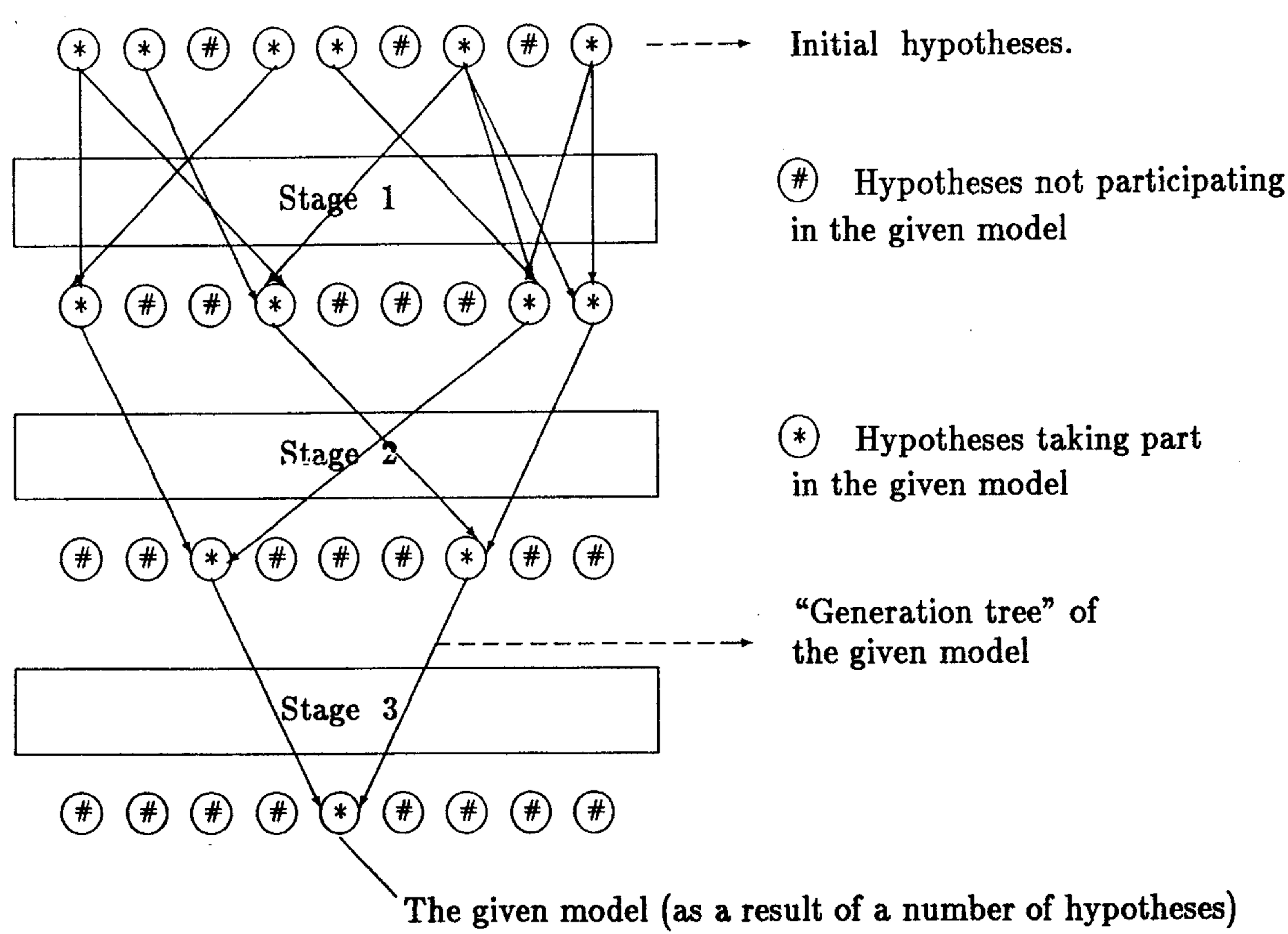
3. In the end, we compute the forecast accuracy measures based on all errors obtained.

External Criteria & Cross Validation

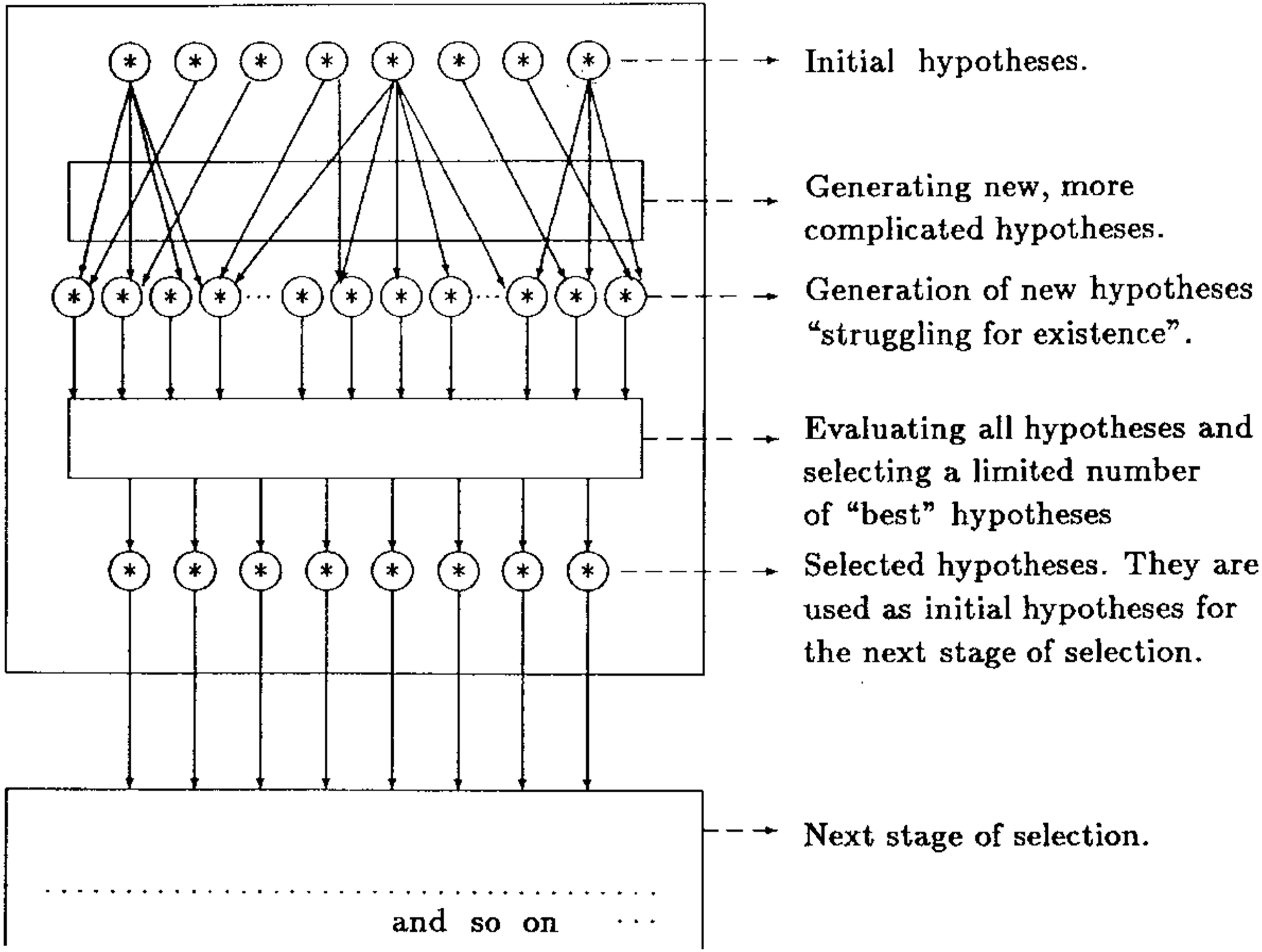
Using a-priori information



Multi-Stage Selection Algorithm



Pair-Wise Selection Using External Criteria

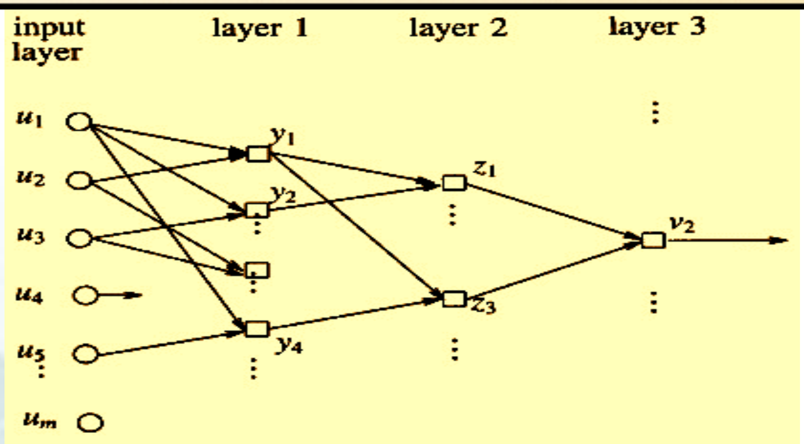


M&M

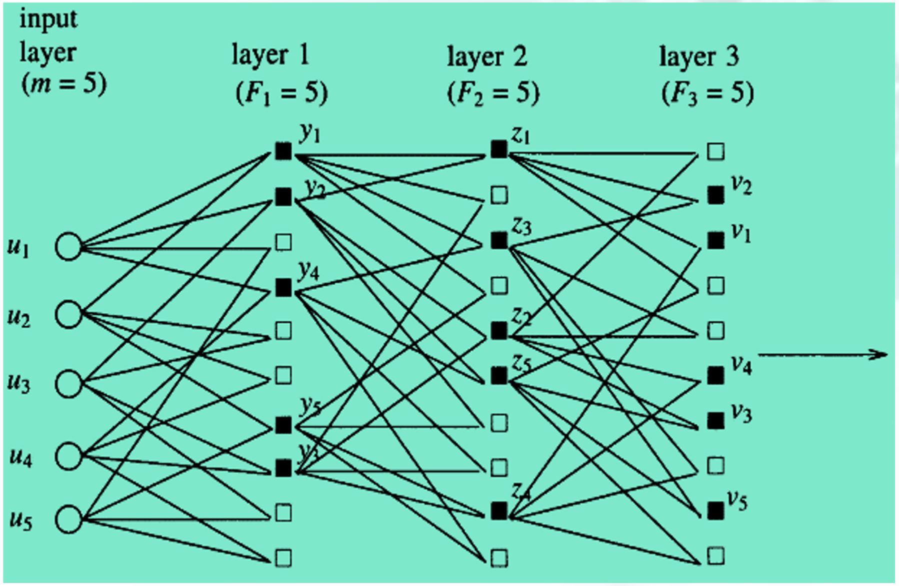
Statistical Learning Networks

Multilayer Net of Active Neurons (MLNAN)

In this approach, neither the number of neurons and the number of layers in the network, nor the actual behavior of each created neuron is predefined. The modeling process is self-organizing because all of them (the number of neurons, the number of layers, and the actual behavior of each created neuron) are adjusting during the process of self-organization.



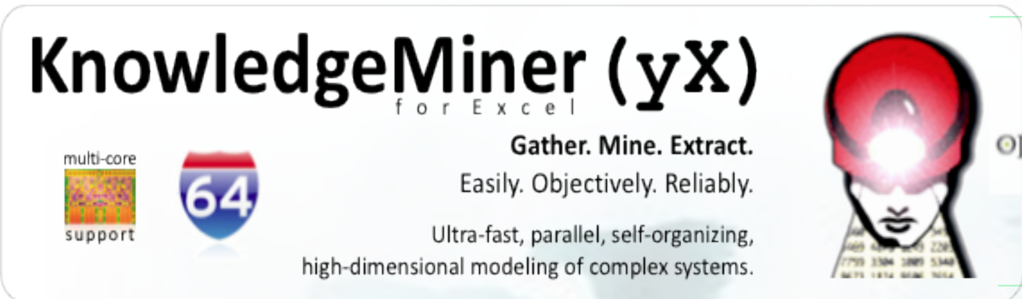
Multilayer network structure representing the output flow to unit 2 of layer 3

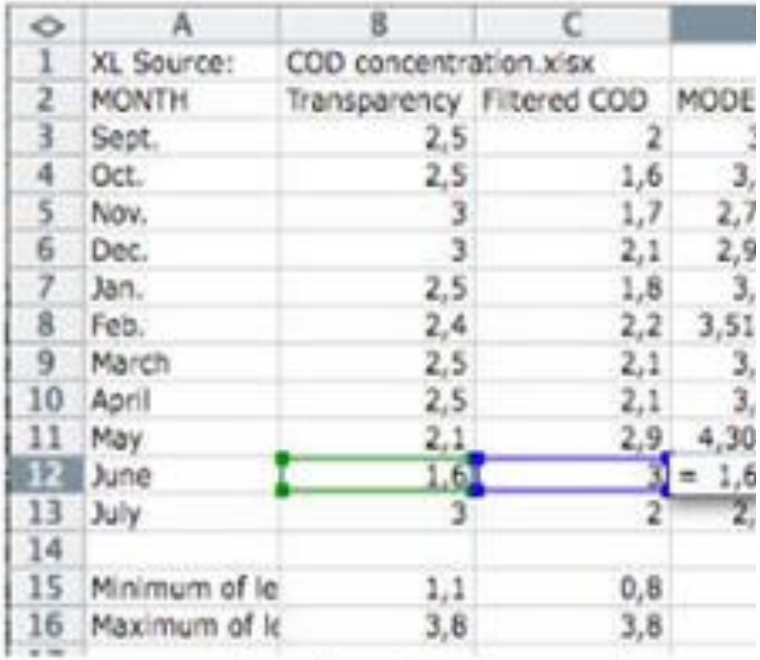


Multilayer network structure with five input arguments and selected nodes:

This method grows a tree-like network out of data of input and output variables in a pair-wise combination and competitive selection from a simple single unit to a desired final solution that does not have a predefined model. The basic idea is that first the elements on a lower level are estimated and the corresponding intermediate outputs are computed and then the parameters of the elements of the next level are estimated.

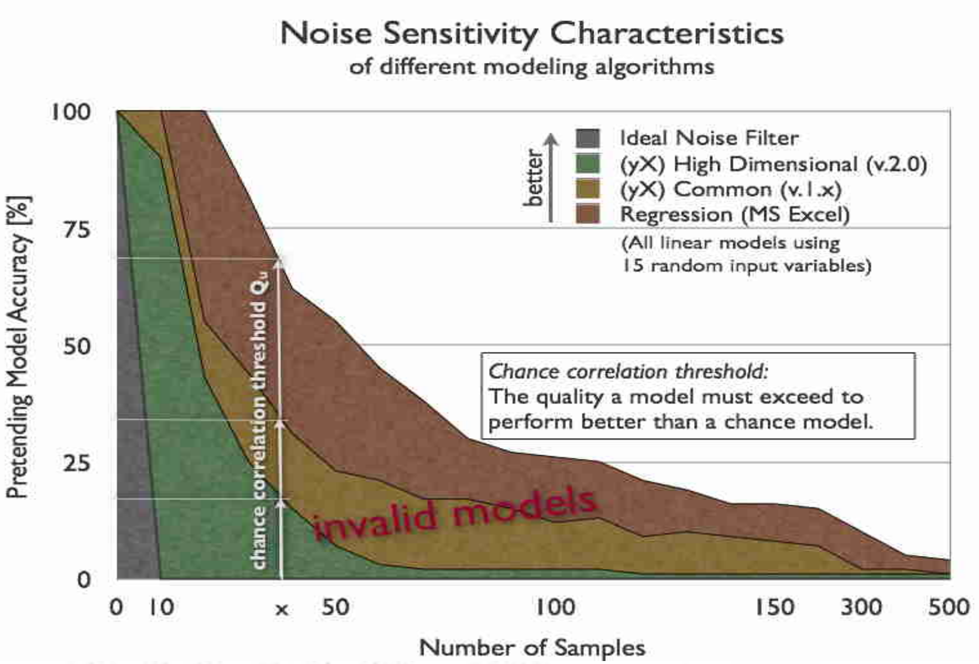
Self-Organizing Data Mining - Applications





	A	B	C	
1	XL Source:	COD concentration.xlsx		
2	MONTH	Transparency	Filtered COD	MODE
3	Sept.	2,5	2	
4	Oct.	2,5	1,6	3,
5	Nov.	3	1,7	2,7
6	Dec.	3	2,1	2,9
7	Jan.	2,5	1,8	3,
8	Feb.	2,4	2,2	3,51
9	March	2,5	2,1	3,
10	April	2,5	2,1	3,
11	May	2,1	2,9	4,30
12	June	1,6	3	= 1,6
13	July	3	2	2,
14				
15	Minimum of le	1,1	0,8	
16	Maximum of le	3,8	3,8	

Analytical model implemented in a new Excel worksheet by (yX) for Excel.



Walla Walla University
SEVENTH-DAY ADVENTIST HIGHER EDUCATION

Who's using it: BOEING, NASA, Mobil, Pfizer

intel Software Partner · Mac · Universal

KnowledgeMiner (yX) for Excel

multi-core support · 64

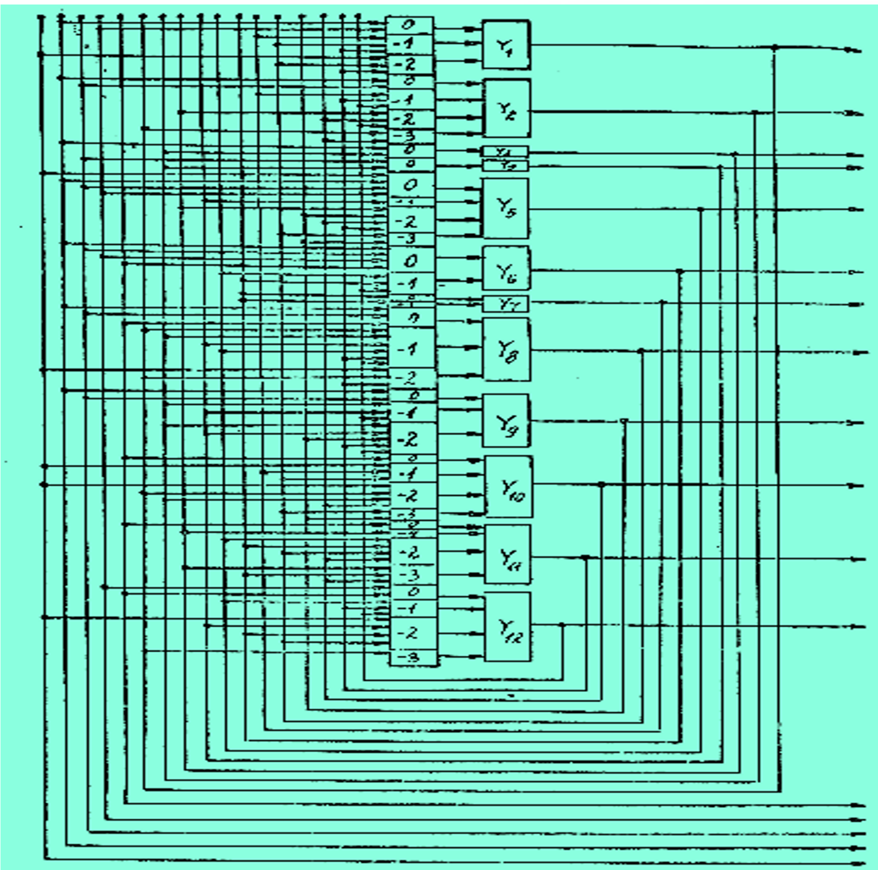
Gather. Mine. Extract.
Easily. Objectively. Reliably.

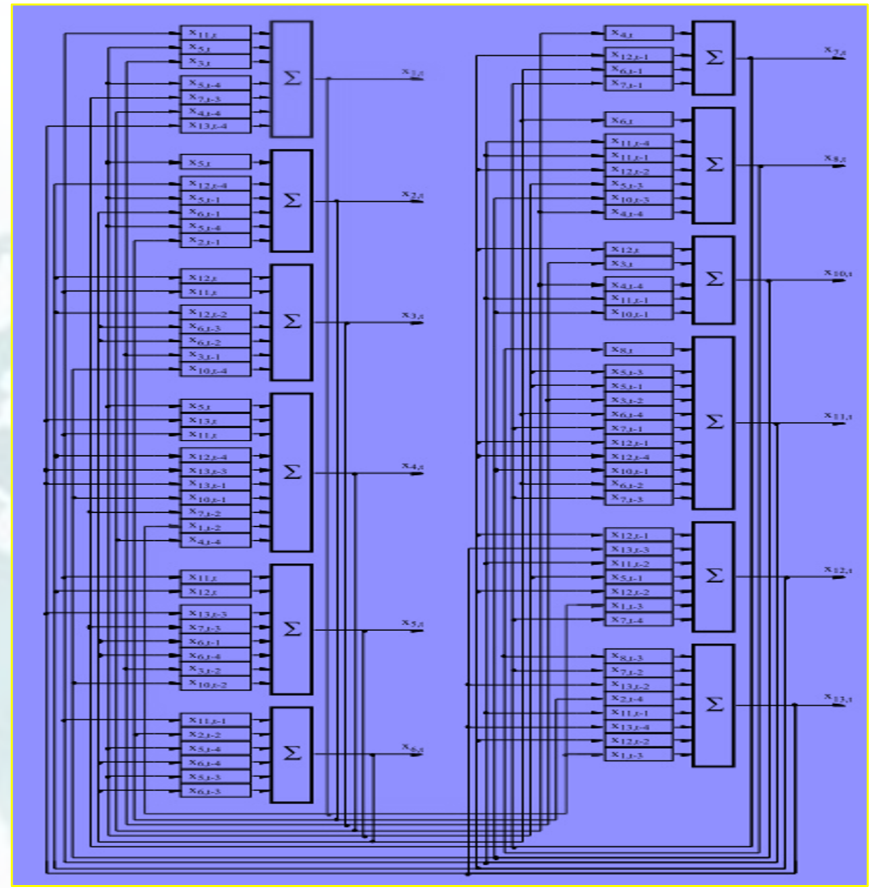
Ultra-fast, parallel, self-organizing, high-dimensional modeling of complex systems.

Berlin, Germany - KnowledgeMiner Software today is pleased to announce the remarkable application of their critically acclaimed collection of self-organizing, automated forecasting technologies implemented in Insights for OS X to a set of high-priority probabilistic energy forecasting problems. Based on a 12-week rolling real-world forecasting scenario in four categories - electric load, electricity price, wind and solar power forecasting - for up to 10 sites, Insights' self-learning forecasting skills ended up in top five positions in all categories of the Global Energy Forecasting Competition (GEFCom2014) organized by the IEEE Power & Energy Society and the University of North Carolina at Charlotte.

(KnowledgeMiner Info info@knowledgeminer.eu)

System's graph of the generated macroeconomic models for Bulgarian and German Economy





Results show that SLNs (MLNAN) are even able to develop complex models reliably and achieve low overall error rates

Walla Walla University
SEVENTH-DAY ADVENTIST HIGHER EDUCATION

SODM with GMDH - Applications

Prediction errors for forecasts calculated using SIMUR II model

Variable Equation	MAPE%				Average for the model:			
	1977	1978	1979	1980	MAPE%	CV(RMSE)%	R	U
1	8.28%	2.58%	4.52%	3.85%	4.81%	5.10%	0.999	0.0260
2	4.26%	6.45%	10.16%	0.61%	5.37%	6.40%	0.975	0.0040
3	3.71%	6.54%	8.31%	0.62%	4.80%	5.84%	0.977	0.0310
4	5.81%	6.15%	15.67%	0.56%	7.05%	8.93%	0.993	0.0078
5	6.17%	0.35%	6.32%	3.81%	4.16%	4.47%	0.998	0.0022
6	0.48%	0.40%	0.04%	1.36%	0.57%	1.82%	0.996	0.0003
7	1.43%	1.15%	1.44%	1.64%	1.42%	1.77%	0.986	0.0035
8	0.96%	0.83%	0.22%	0.62%	0.66%	0.72%	0.995	0.0001
9	5.24%	6.15%	5.60%	8.23%	6.31%	6.68%	0.995	0.0044
10	0.77%	2.22%	1.20%	3.34%	1.88%	2.31%	0.998	0.0005
11	13.89%	6.26%	6.39%	3.86%	7.60%	7.72%	0.996	0.0056
12	1.03%	3.16%	3.16%	1.18%	2.13%	3.28%	0.995	0.0010
Average	4.34%	3.52%	5.25%	2.47%	3.90%	4.59%	0.992	0.0098

SODM with GMDH - Applications

Prediction errors for German economy forecasts using *Knowledge-Miner*

Variable	% Differences Between Predictions & Real Data			MAPE	MSE (%)
	1988	1989	Average %		
Y_{1t}	2.58%	3.85%	3.22%	3.48%	5.10%
Y_{2t}	6.45%	0.61%	3.53%	3.73%	6.40%
Y_{3t}	6.54%	0.62%	3.58%	3.84%	5.85%
Y_{4t}	6.15%	0.56%	3.36%	3.59%	8.93%
Y_{5t}	0.35%	3.81%	2.08%	2.30%	4.47%
Y_{6t}	0.04%	1.36%	0.70%	1.10%	1.82%
Y_{7t}	16.51%	17.60%	17.06%	17.80%	10.12%
Y_{8t}	0.83%	0.62%	0.73%	1.15%	0.72%
Y_{9t}	6.15%	8.23%	7.19%	7.43%	6.68%
Y_{10t}	2.22%	3.34%	2.78%	2.98%	2.31%
Y_{11t}	6.26%	3.86%	5.06%	5.45%	7.72%
Y_{12t}	3.16%	1.18%	2.17%	2.48%	3.28%
Mean	4.77%	3.80%	4.29%	4.61%	5.28%

Data Mining and Predictive Modelling

The family of models "S I M U R"

Model No . Year of design	Objectives – Improving Accuracy	Characteristics
SIMUR 0 (1977—1978)	First step of the SIMUR project. Traditional methods used. Accuracy = 14%.	A one—product macroeconomic model in the form of SSE with 5 equations. Contains 5 endogenous, 5 lag and 1 exogenous variables.
SIMUR I (1978—1980)	Analysis of possibilities for automatic synthesis of SSE during the run of multi-stage selection procedure. Accuracy = 2.7%.	A one—product macroeconomic model in the form of SSE with 5 equations. Contains the same set of variables.
SIMUR II (1981—1982)	Design and experimenting of a programming system for automatic holding of simulation experiments with SSE. Analysis of different criteria for the precision's estimating of SSE. Accuracy = 2.0%.	Aggregated macroeconomic model in the form of 12 interdepending simultaneous equations. Contains 12 endogenous, 5 exogenous and 26 lag variables with lag of up to 3 years.
SIMUR III (1983—1985)	Improving the MSSP for synthesis of SSE with many equations. Simulating and forecasting of main macroeconomic indexes. Accuracy < 1%.	Macroeconomic simulation model. Contains 39 equations and 39 endogenous, 7 exogenous and 82 lag variables with a time lag of up to 5 years.
SIMUR IV (1989 -)	Next step of the SIMUR project.	Multisector macroeconomic model. Contains more than 100 equations.

Model-Based Business Games

Walla Walla University
SEVENTH-DAY ADVENTIST HIGHER EDUCATION

About WWU | Academics | Life at WWU | Resources | Attend WWU | Email | MyWWU | Login

HOME > ABOUT WWU > GENERAL INFORMATION > NEWS >

Walla Walla University

NAVIGATION

News
Archive
Name Change

RELATED LINKS

Inside WWU Newsletter

WHAT'S HAPPENING

APIC Chinese New Year Party ...

Quicklinks

Business Professor Designs Game to Help Industry Professionals

Motzev Has Shared Research Results at Worldwide Conferences

By: Becky St. Clair

Mihail Motzev, School of Business professor

Who says professionals can't have fun? Mihail Motzev, professor in the School of Business at Walla Walla University, spent three years developing what is essentially a game for businesspeople. His latest project, titled "Intelligent Techniques in Simulations and Management Games–A Hybrid Approach: Multi-Stage Selection Procedures for Complex Systems Model Building" was nominated and approved by WWU faculty for a faculty research grant three consecutive years.

"It's one of my favorite areas of research," says Motzev. "I enjoy the work and the presentation as much as I enjoy the end result."

As a member of the International Simulation and Gaming Association (ISAGA), Motzev has presented results from his research at many conventions, most recently in Romania, Poland, and Spokane, Wash. He has also been invited to present at the ISAGA/IFIP (International Federation for Information Processing) world conference in Sweden this summer.

Model-Based Business Games

Students compete in teams in a computerized business simulation to see who can best manage the resources of a corporation to maximize profitability and minimize total cost:

Source: ISAGA 2011

New Product – An Integrated Simulation Game In Business Training & Education

Strengths & Advantages in Simulation

SLNs (MLNAN) provide opportunities for:

- shorten the time;
- reducing the cost;
- reducing the efforts in model building;
- increasing the ***accuracy*** of the model;
- developing even complex models reliably with low overall error rates.

Walla Walla University
SEVENTH-DAY ADVENTIST HIGHER EDUCATION

Applications

Example: *Model-Based Business Games*

Original Version	New Version
A one-product macro-economic model developed as a system of five SE. Contains five endogenous, one exogenous, and five lag variables.	A one—product macroeconomic model with the same structure. Contains same set of variables.
Indirect OLS used to estimate unknown coefficients in equations.	Model synthesized using the hybrid algorithm.
Model accuracy - mean squared error (MSE) = 14%	Model accuracy MSE = 2.7%

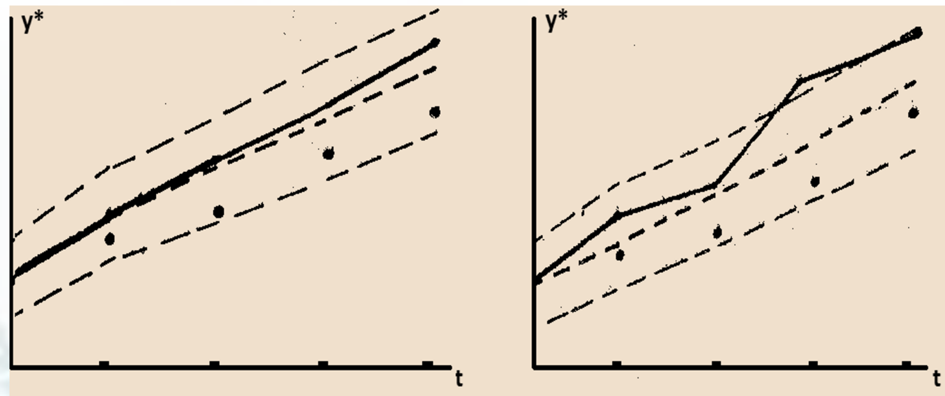
Data Mining and Predictive Modelling

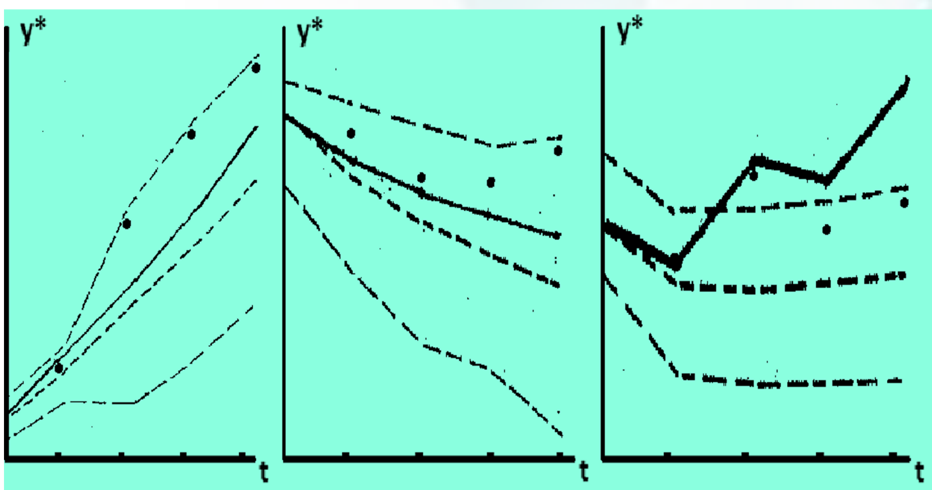
Improving Model Accuracy

Statistics and test results for ***ex-post*** forecasts

Eq. no	Coefficient of multiple correlation (R)	Coefficient of multiple determination (R^2)	F value	CV(RMSE)%	MAPE %
1	0.9949	0.9898	355.81	3.40%	2.69%
2	0.9980	0.9960	936.50	2.16%	1.77%
3	0.9991	0.9982	1848.52	1.76%	1.40%
4	0.9637	0.9287	47.69	0.66%	0.52%
5	0.9522	0.9067	24.30	11.09%	7.13%
Total	0.9816	0.9635	642.56	3.81%	2.70%

Ex-ante **predictions four years ahead for variables $y_{1,t}$ and $y_{2,t}$**





Ex-ante predictions 4 years ahead for variables $y_{3,t}$, $y_{4,t}$ and $y_{5,t}$

Eq. no	Coefficient of multiple correlation (R)	Coefficient of multiple determination (R^2)	F value	CV(RMSE) %	MAPE %
1	0.9928	0.9857	254.55	4.03%	3.57%
2	0.9925	0.9851	242.42	4.22%	3.78%
3	0.9993	0.9986	1664.17	3.84%	3.39%
4	0.8288	0.6869	8.05	1.39%	1.11%
5	0.9362	0.8765	17.76	12.76%	8.56%
Total	0.9499	0.9023	437.39	5.25%	4.08%

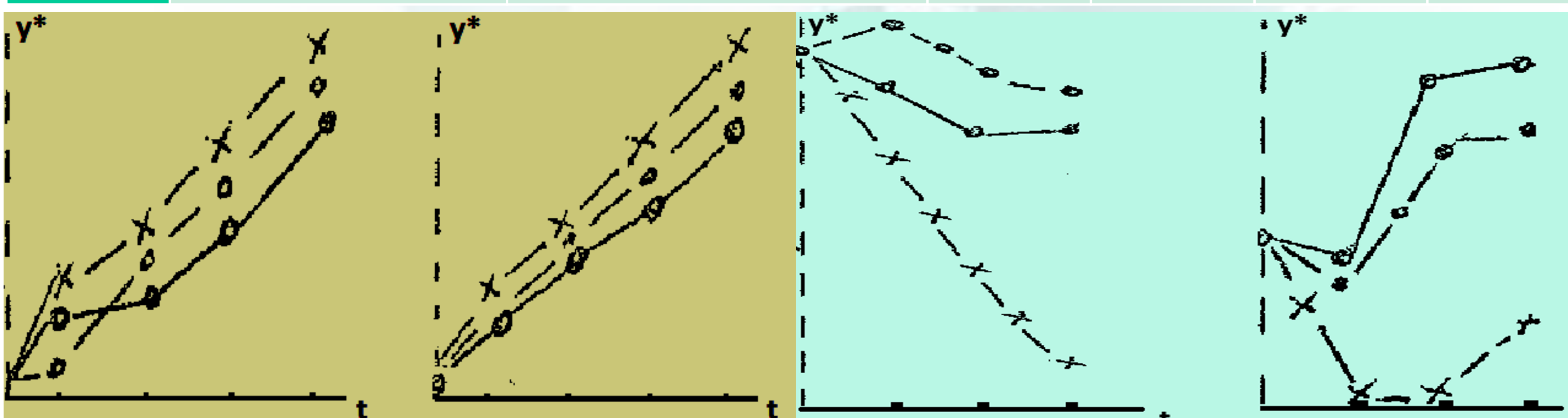
Statistics and test results for ***ex-ante*** forecasts

Data Mining and Predictive Modelling

Improving Model Accuracy

Comparisons for *ex-post* forecasts from stationary and non-stationary models

Equation No.	Coefficient of multiple determination (R^2)		von Neumann test for Autocorrelation (Q)		CV(RMSE)%		MAPE%	
	Stationary	Non-Stat.	Stationary	Non-Stat.	Stationary	Non-Stat.	Stationary	Non-Stat.
1	0.9898	0.9920	1.010	0.965	3.40%	2.78%	2.69%	2.35%
2	0.9960	0.9980	0.735	1.212	2.16%	1.44%	1.77%	1.23%
3	0.9982	0.9980	1.070	1.069	1.76%	1.63%	1.40%	1.34%
4	0.9287	0.9900	0.846	1.081	0.66%	0.81%	0.52%	0.72%
5	0.9067	0.9460	1.020	1.233	11.09%	7.61%	7.13%	6.73%
For the model	xxx		xxx		3.81%	2.85%	2.70%	2.48%



Real values Non-stationary

Stationary

Ex-ante **prediction comparisons for three years ahead with stationary and non-stationary models**

Walla Walla University
SEVENTH-DAY ADVENTIST HIGHER EDUCATION

Prediction Accuracy

Experimental Test Results - 2018

Best Model	Second Best	Third Best
MLNAN: MASE: 0.0414 MPE = 1.42% MAPE = 1.42% CV(RMSE) = 1.56%	Triple Exponential MASE = 0.0627 MPE = -0.57% MAPE = 1.76% CV(RMSE) = 2.45%	Multiple Autoregression MASE = 0.0908 MPE = 2.03% MAPE = 2.58% CV(RMSE) = 3.17%

Prediction Accuracy

Experimental Test Results - 2019

Best Model	Second Best	Third Best
MLNAN: MASE: 0.0446 MPE = 1.55% MAPE = 1.55% CV(RMSE) = 1.56%	Multiple Regression with Time and Dummy Seasonal MASE = 0.0508 MPE = -1.09% MAPE = 1.59% CV(RMSE) = 1.56%	Triple Exponential MASE = 0.0627 MPE = -0.57% MAPE = 1.76% CV(RMSE) = 2.45%

Walla Walla University
SEVENTH-DAY ADVENTIST HIGHER EDUCATION

Thank You!

and I'll

See You again...