

# Artificial Intelligence Models for Knowledge Producing

Jelenka Savkovic-Stevanovic

University of Belgrade, Faculty of Technology and Metallurgy

DOI: <https://doi.org/10.51584/IJRIAS.2025.10020054>

Received: 11 February 2025; Accepted: 15 February 2025; Published: 21 March 2025

## ABSTRACT

In this paper artificial intelligence models are structured and discussed. Generative models vs. conditional models, regression models vs. classification models, inductive models vs. deductive models and algorithms were studied. Complex model with various levels was expressed by hierarchy equation. Prediction the conditional probability was discussed and described. In this paper for the complex model in various level the hierarchy equation was derived. The first time in literature, this paper acknowledges the significance of ethics in artificial intelligence making.

**Keywords:** Artificial intelligence, models, algorithms, hypothesis, expression,

## INTRODUCTION

There are two approaches at complex system modelling. The first approach is identified with structured knowledge following deduced reasoning, that is approach which deduced relations on that problem from existing theory. The second approach is identified with post empirical knowledge following inductive approach in which model developing from sampling data. These two approaches represent complementary stages of system modelling.

In science models have the first of all two purposes logical enabling some conclusions to which can not be come by the other methods and epistemological to express knowledge and to enable to extent knowledge about real system. Modelling methods linking previous knowledge, known facts, scientific laws and hypothesis so that from them can conclude new knowledge about given domain [1]-[7].

In general model is any phenomenon, which similar to studying phenomenon according to definition, and two systems are structural similar each other, only if exists isomorphism between them. Expression model is multi-significant and has significance interpretation of the set values formulas, whose all terms are adequate and significance symbolic representation of system.

Structured system models enabling parameters analysis on all levels structural description. By changing of these parameters changes system characteristics with the aim seeking out optimal conditions, and over them is introduced new materials in the system, that enabling innovation development. By deriving the new relations in the model, it is introduced the new operation on the system, that producing innovation, too.

Non-structured system models are not taking into account inner system structure, already only changes of the system. At the structured models, the inner structure and their changes taking into account. Non-structured models, sometimes are signed as models with parameters group.

An artificial intelligence (AI) model is the output of the algorithm which applied in an dataset.

In this paper AI models were defined, structured, classified, and illustrated relations between them.

## What An Ai Model Means?

An AI model is a program that has been trained on a set of data to recognize certain patterns or make certain decisions without further human intervention. Artificial intelligence models apply different algorithms to relevant data inputs to achieve the tasks, or output, they've been programmed for.

Simply put, an AI model is defined by its ability to autonomously make decisions or predictions, rather than simulate human intelligence. Among the first successful AI models were checkers- and chess-playing programs in the early 1950s: the models enabled the programs to make moves in direct response to the human opponent, rather than follow a pre-scripted series of moves.

Different types of AI models are better suited for specific tasks, or *domains*, for which their particular decision-making logic is most useful or relevant. Complex systems often employ multiple models simultaneously, using ensemble learning techniques like bagging, boosting or stacking.

$$M_1, M_2, M_3, \dots, M_n \in M_{n+1} \quad (1)$$

As AI tools grow increasingly complex and versatile, they require increasingly challenging amounts of data and computing power to train and execute. In response, systems designed to execute specific tasks in a single domain are giving way to *foundation models*, pre-trained on large, unlabeled datasets and capable of a wide array of applications. These versatile foundation models can then be fine-tuned for specific tasks.

## Algorithms and Models

Though the two terms are often used interchangeably in this context, they do not mean quite the same thing.

- *Algorithms* are logical procedures, often described in mathematical language or pseudocode, to be applied to a dataset to achieve a certain function or purpose.
- *Models* are the output of an algorithm that has been applied to a dataset.

In simple terms, an AI model is used to make predictions or decisions and an algorithm is the logic by which that AI model operates.

It needs to choose the right AI model for the use case .

Bigger isn't always better when it comes to AI models. Learn how to find the right fit for your business needs. Then get the guide to help you take action.

## Ai Models and Machine Learning

AI models can automate decision-making, but only models capable of machine learning (ML) are able to autonomously optimize their performance over time.

While all ML models are AI, not all AI involves ML. The most elementary AI models are a series of if-then-else statements, with rules programmed explicitly by a data scientist. Such models are alternatively called *rules engines*, *expert systems*, *knowledge graphs* or *symbolic AI*.

Machine learning models use *statistical AI* rather than symbolic AI. Whereas rule-based AI models must be explicitly programmed, ML models are “trained” by applying their mathematical frameworks to a sample dataset whose data points serve as the basis for the model’s future real-world predictions.

Machine learning model techniques can generally be separated into three broad categories: *supervised learning*, *unsupervised learning* and *reinforcement learning*.

- *Supervised learning*: also known as “classic” machine learning, supervised learning requires a human expert to label training data. A data scientist training an image recognition model to recognize dogs and cats must label sample images as “dog” or “cat”, as well as key features-like size, shape or for-that inform those primary labels. The model can then, during training, use these labels to infer the visual characteristics typical of “dog” and “cat”.

- **Unsupervised learning:** Unlike supervised learning techniques, unsupervised learning does not assume the external existence of “right” or “wrong” answers, and thus does not require labeling. These algorithms detect inherent patterns in datasets to cluster data points into groups and inform predictions. For example, e-commerce businesses like Amazon use unsupervised association models to power recommendation engines.
- **Reinforcement learning:** in reinforcement learning, a model learns holistically by trial and error through the systematic rewarding of correct output (or penalization of incorrect output). Reinforcement models are used to inform social media suggestions, algorithmic stock trading, and even self-driving cars.

*Deep learning* is a further evolved subset of unsupervised learning whose structure of neural networks attempts to mimic that of the human brain. Multiple layers of interconnected nodes progressively ingest data, extract key features, identify relationships and refine decisions in a process called forward propagation [8]-[14]. Another process called backpropagation applies models that calculate errors and adjust the system's weights and biases accordingly. Most advanced AI applications, like the large language models (LLMs) powering modern chat-bots, utilize deep learning. It requires tremendous computational resources [8]-[14].

### Deductive Vs. Inductive Models

An deductive model learning systems includes knowledge based system, design, operation and optimization algorithms. System application for various examples can be expression.

At the inductive model the datasets in a given space integrated and analyzed decision support system. These models involve human reasoning and supervision.

### Generative Vs. Discriminative Models

One way to differentiate machine learning models is by their fundamental methodology: most can be categorized as either *generative* or *discriminative*. The distinction lies in how they model the data in a given space.

*Generative* algorithms, which usually entail unsupervised learning, model the *distribution* of data points, aiming to predict the *joint probability*  $P(x,y)$  of a given data point appearing in a particular space. A generative computer vision model might thereby identify correlations like “things that look like cars usually have four wheels” or “eyes are unlikely to appear above eyebrows.”

These predictions can inform the generation of outputs the model deems highly probable. For example, a generative model trained on text data can power spelling and autocomplete suggestions; at the most complex level, it can generate entirely new text. Essentially, when a large language model outputs text, it has computed a high probability of that sequence of words being assembled in response to the prompt it was given.

Other common use cases for generative models include image synthesis, music composition, style transfer and language translation.

Examples of generative models include:

- **Diffusion models:** diffusion models gradually add Gaussian noise to training data until it's unrecognizable, then learn a reversed “denoising” process that can synthesize output (usually images) from random seed noise [8].
- **Variational autoencoders (VAEs):** VAEs consist of an encoder that compresses input data and a decoder that learns to reverse the process and map likely data distribution.

- **Transformer models:** Transformer models use mathematical techniques called “attention” or “self-attention” to identify how different elements in a series of data influence one another. The “GPT” in Open AI’s Chat-GPT stands for “Generative Pretrained Transformer.”

**Conditional** algorithms, which usually entail supervised learning, model the boundaries between classes of data (or “decision boundaries”), aiming to predict the *conditional probability*  $P(y|x)$  of a given data point ( $x$ ) falling into a certain class ( $y$ ). For example, a discriminative computer vision model might learn the difference between “car” and “not car” by discerning a few key differences (like “if it doesn’t have wheels, it’s not a car”), allowing it to ignore many correlations that a generative model must account for. Discriminative models thus tend to require less computing power.

Prediction by conditional probability can express as:

Given an event  $x$  with nonzero probability,

$$P(x) > 0 \quad (2)$$

it can define the conditional probability of  $y$  assuming  $x$ , by

$$P(y|x) = \frac{P(yx)}{P(x)} \quad (3)$$

In words  $P(y|x) > 0$  equals to probability of the event  $yx$ , the part of  $y$  included in  $x$ , divided by the probability of  $x$ . Clearly, if  $y$  and  $x$  have no common elements (mutual, exclusively) then  $P(y|x) = 0$ . If recall from eq.(2) that, with  $x$  an event such that

$$P(y|x) \neq 0. \quad (4)$$

Discriminative models are, naturally, well suited to classification tasks like sentiment analysis -but they have many uses for making the ethical decision. For example, decision tree and random forest models break down complex decision-making processes into a series of nodes, at which each “leaf” represents a potential classification decision.

While discriminative or generative models may *generally outperform* one another for certain real -world use cases, many tasks could be achieved with either type of model. For example, discriminative models have many uses in natural language processing (NLP) and often outperform generative AI for tasks like machine translation (which entails the generation of translated text).

Similarly, generative models can be used for classification using Bayes’ theorem [1],[15]. Rather than determining which side of a decision boundary an instance is on (like a discriminative model would), a generative model could determine the probability of each class generating the instance and pick the one with higher probability.

Many AI systems employ both in tandem. In a generative adversarial network, for example, a generative model generates sample data and a discriminative model determines whether that data is “real” or “fake.” Output from the discriminative model is used to train the generative model until the discriminator can no longer discern “fake” generated data [16].

## A case study: Diagnostic AI model

Let consider the input data of a plant in Fig. 1 [1]. A systematic cause-consequence analysis gives the results which are summarized in the form of a fault tree model. It follows the structure of a generic fault tree point to

the release of materials, of an event tree from this point to the impact of the release on people, the plant and the environment.

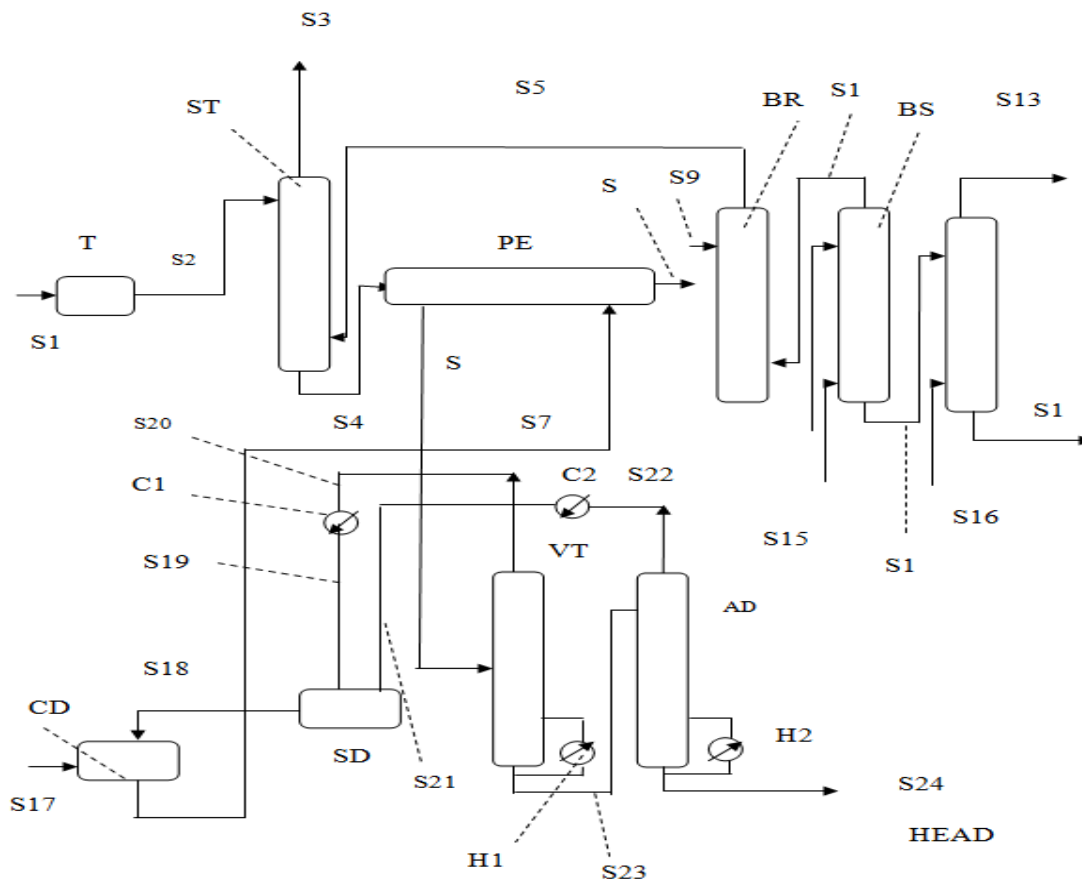


Fig.1 The plant input data.

Equipment state are described in qualitative term such as: *closed*, *open*, *failed*, *blocked* and *leak*. The following block are considered *blockage*, *leakage*, *malfunction* or *miss operation*. The study of fault detection and diagnostic is concerned with designing system that can assist the human operator detecting and diagnosing equipment faults in order to prevent accidents.

Its knowledge base composed of equipment and material streams and data base of occurred symptoms and faults at a single units. M, B, and L are independent Boolean variables representing the basic events: malfunction, blockage, leakage, respectively.

The independent variables should be replaced by relative frequencies of the events  $p_i$ . The Boolean operators *AND* or *OR* could be replaced by the algebraic operators  $AND(p_1, p_2, p_3, \dots, p_n)$  and  $OR(p_1, p_2, p_3, \dots, p_n)$  producing the output frequency  $p_y$  from the input frequencies  $p_1, p_2, p_3, \dots, p_n$ .

For the quantitative model term *relative frequency* was used instead of *event probability*.

Probabilistic variables must fulfill this:

$$\overline{p_i} = 1 - p_i \quad (5)$$

The  $AND(p_1, p_2, p_3, \dots, p_n)$  operator assigns for  $p_y$  the value  $n$ .

$$p_y = \prod_{i=1}^n p_i \quad (6)$$

Analogously for the  $OR(p_1, p_2, p_3, \dots, p_n)$  operator:

$$p_y = \sum_{i=1}^n p_i \quad (7)$$

Equation (5) does not fulfill requirement that the relative frequencies must lie in the range  $0 \leq p_i \leq 1$ . Therefore, equation (6) is transformed using Morgan's rule:

$$p_1 + p_2 + p_3 + \dots + p_n = \overline{p_1 \bullet p_2 \bullet p_3 \bullet \dots \bullet p_n} \quad (8)$$

can express in the form

$$p_y = 1 - \prod_{i=1}^n (1 - p_i)$$

and induced events represent as follows:

$$p_{HEAD} = 1 - p_{S24}$$

$$p_{S24} = 1 - (1 - p_{B(S24)})(1 - p_{L(S24)})(1 - p_{AD})$$

$$p_{AD} = 1 - (1 - p_{B(AD)})(1 - p_{M(AD)})(1 - p_{H2})(1 - p_{S22})(1 - p_{S23})$$

$$p_{S22} = 1 - (1 - p_{B(S22)})(1 - p_{L(S22)})(1 - p_{C2})$$

$$p_{H2} = 1 - (1 - p_{B(H2)})(1 - p_{L(H2)})(1 - p_{M(H2)})$$

$$p_{C2} = 1 - (1 - p_{B(C2)})(1 - p_{L(C2)})(1 - p_{M(C2)})(1 - p_{S21})$$

$$p_{VT} = 1 - (1 - p_{B(VT)})(1 - p_{L(VT)})(1 - p_{M(VT)})(1 - p_{H1})(1 - p_{S20})(1 - p_{S6})$$

$$p_{S21} = 1 - (1 - p_{B(S21)})(1 - p_{L(S21)})(1 - p_{SD})$$

$$p_{SD} = 1 - (1 - p_{B(SD)})(1 - p_{L(SD)})(1 - p_{M(SD)})(1 - p_{S22})(1 - p_{S19})$$

$$p_{H1} = 1 - (1 - p_{B(H1)})(1 - p_{L(H1)})(1 - p_{M(H1)})$$

$$p_{S20} = 1 - (1 - p_{B(S20)})(1 - p_{L(S20)})$$

$$p_{S18} = 1 - (1 - p_{B(S18)})(1 - p_{L(S18)})$$

$$p_{S19} = 1 - (1 - p_{B(S19)})(1 - p_{L(S19)})(1 - p_{C1})$$

$$p_{C1} = 1 - (1 - p_{B(C1)})(1 - p_{L(C1)})(1 - p_{M(C1)})(1 - p_{S20})$$

$$p_{S6} = 1 - (1 - p_{B(S6)})(1 - p_{L(S6)})(1 - p_{PE})$$

$$p_{PE} = 1 - (1 - p_{B(PE)})(1 - p_{L(PE)})(1 - p_{S7})(1 - p_{S4})$$

$$p_{S7} = 1 - (1 - p_{B(S7)})(1 - p_{L(S7)})(1 - p_{CD})$$

$$p_{CD} = 1 - (1 - p_{B(CD)})(1 - p_{L(CD)})(1 - p_{M(CD)})(1 - p_{S17})(1 - p_{S18})$$



$$pS17 = 1 - (1 - pB(S17))(1 - pL(S17))(1 - pNOT\_SUPPLY(S17))$$

$$pS4 = 1 - (1 - pB(S4))(1 - pL(S4))(1 - pST)$$

$$pST = 1 - (1 - pB(ST))(1 - pL(ST))(1 - pS5)(1 - pS3)(1 - pS2)$$

$$pS2 = 1 - (1 - pB(S2))(1 - pL(S2))(1 - pTD)$$

.  
.  
.  
etc.

## Regression Vs. Classification Models

Another way to categorize models is by the nature of the tasks they are used for. Most classic AI model algorithms perform either *classification* or *regression*. Some are suitable for both, and most foundation models leverage both kinds of functions.

This terminology can, at times, be confusing. For example, logistic regression is a discriminative model used for classification.

Regression models predict *continuous* values (like price, age, size or time). They're primarily used to determine the relationship between one or more independent variables ( $x$ ) and a dependent variable ( $y$ ): *given  $x$ , predict the value of  $y$ .*

- *Algorithms* like linear regression, and related variants like quantity regression, are useful for tasks like forecasting, analyze pricing elasticity, and assessing risk.
- *Algorithms* like polynomial regression and support vector regression (SVR) model complex non-linear relationships between variables.
- *Certain generative models*, like autoregression and variational autoencoders, account for not only correlative relationships between past and future values, but also *causal* relationships. This makes them particularly useful for forecasting weather scenarios and predicting extreme climate events.

Classification models predict *discrete* values. As such, they're primarily used to determine an appropriate label or to categorize (i.e., classify). This can be a *binary* classification-like "yes or no," "accept or reject"- or a *multi-class* classification (like a recommendation engine that suggests Product A, B, C or D).

Classification algorithms find a wide array of uses, from straightforward categorization to automating feature extractions in deep learning networks to healthcare advancements like diagnostic image classification in radiology.

Common examples include:

- Naive bayes: a generative supervised learning algorithm commonly used in spam filtering and document classification.
- Linear discriminant analysis: used to resolve contradictory overlap between multiple features that impact classification.
- Logistic regression: predicts continuous probabilities that are then used as proxy for classification range.

## Ai Models Training

The “learning” in machine learning is achieved by training models on sample datasets. Probabilistic trends and correlations discerned in those sample datasets are then applied to performance of the system’s function.

In supervised and semi-supervised learning, this training data must be thoughtfully labeled by data scientists to optimize results. Given proper feature extraction, supervised learning requires a lower quantity of training data overall than unsupervised learning.

Ideally, ML models are trained on real-world data. This, intuitively, best ensures that the model reflects the real-world circumstances that it’s designed to analyze or replicate. But relying solely on real-world data is not always possible, practical or optimal.

The more parameters a model has, the more data is needed to train it. As deep learning models grow in size, acquiring this data becomes increasingly difficult. This is particularly evident in LLMs: both Open-AI’s GPT-3 and the open source BLOOM have over 175 billion parameters.

Despite its convenience, using publicly available data can present regulatory issues, like when the data must be anonymized, as well as practical issues. For example, language models trained on social media threads may “learn” habits or inaccuracies not ideal for enterprise use.

*Synthetic data* offers an alternative solution: a smaller set of real data is used to generate training data that closely resembles the original and eschews privacy concerns.

## DISCUSSION

1. Complex models can divided in various levels and put into hierarchy model. The hierarchy equation was derived.
- 2 :Diffusion models gradually add Gaussian noise to training data until it’s unrecognizable.
3. Conditional probability prediction for discriminatory models was expressed.
4. Certain generative models are represented not only correlative relationships between past and future values, but also causal relationships.
5. Classification algorithms find a wide array of uses, in deep learning networks.
6. Beyond summarizing AI models structure the relations between them have given in Fig. 2.

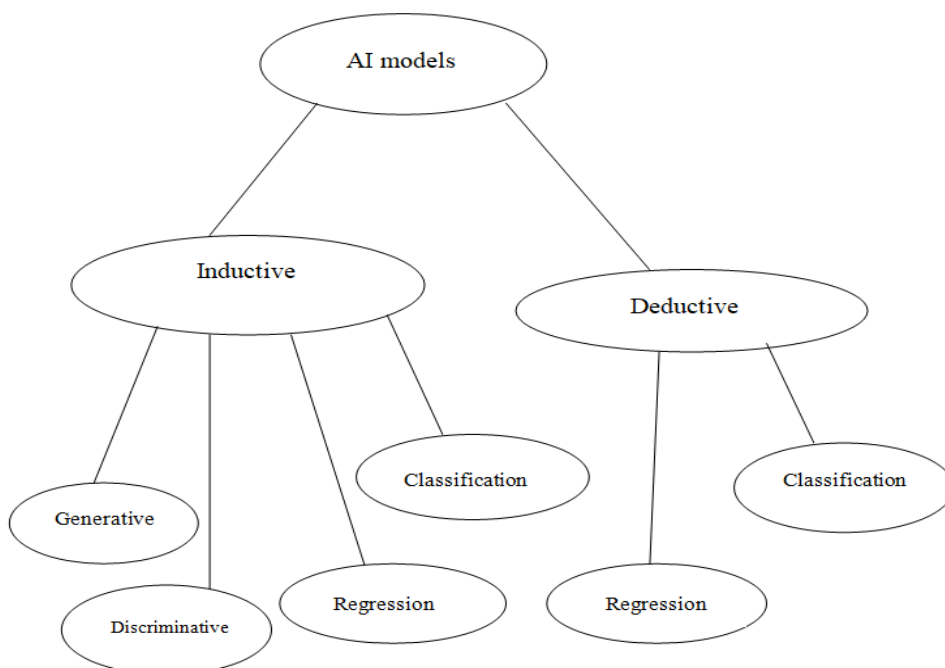


Fig. 2 The graph



## CONCLUSION

In this paper distributed modelling, perception and recognition by AI intelligence models were studied. Classification and structured models were examined. Models dealing with mental processes of decision makers is a part of cognitive science. It will become increasingly important for model builder and users to have a clear and strong code of ethics to guide them in making the ethical decision they surely will have to face.

In this paper the equation for multilevel complex models hierarchy was derived. For discriminative models prediction the conditional probability was expressed.

The main contributions of this paper are in general model definition as any phenomenon, which structural similar each other, as well as summarizing AI models structures.

## Notation

*M* - model

$P(y|x)$  - conditional probability

## Abbreviation

*AI* - artificial intelligence

*LLMs* - large language models (LLMs)

*ML* - machine learning

*VAEs* - Variational autoencoders

*NLP* - natural language processing

*SVR* - support vector regression

## REFERENCES

1. Savkovic Stevanovic J. (2022) Artificial intelligence in chemical engineering practice, LAP-Lambert Academic Publishing, www.lap-publishing.com, ISBN 978-6-20551509-9, www.amazon.com.
2. Sample I., Ian (2017) Google's Deep Mind makes AI program that can learn like a human". The Guardian. Archived from the original on 26 April 2018. Retrieved 26 April 2018.
3. Habibi, Aghdam, Hamed (2017) Guide to convolutional neural networks: a practical application to traffic-sign detection and classification, Heravi, Elnaz Jahani. Cham, Switzerland, , ISBN 9783319575490. OCLC 987790957.
4. Tang A. et al. (2018) Canadian association of radiologists white paper on artificial intelligence in radiology. Can. Assoc. Radiol., J. Assoc. Can. Radiol. 69, 120–135. [PubMed] [Google Scholar].
5. Savkovic Stevanovic J. (2008) Process Engineering intelligent systems, Srbisim, Belgrade, Serbia, ISBN 978-86-911011-1-4.
6. Savkovic-Stevanovic (1995) Process Modelling and Simulation, Faculty of Technology and Metallurgy, Belgrade University, Serbia.
7. Savkovic-Stevanovic J. (2007) Informatics, Faculty of Technology and Metallurgy, Belgrade University, Serbia, ISBN 978-86-7401-244-4.
8. An P. E., M. Brown, Harris C. J. (1995) A global gradient noise covariance expression for stationary real Gaussian inputs, IEEE Trans. Neural Networks, 6, 1549-1551.
9. Jordan M. I., Jacobs R. A. (1994) Hierarchical mixtures of experts and the EM algorithm, Neural computation, 6, 181-214.

10. Savkovic Stevanovic J. (1993) A neural-network model for analysis and optimization of processes, *Computers&Chemical Engineering*, **17**, 411.
11. Savkovic Stevanovic J. (1994) An qualitative model for estimation of plant behavior *Computers&Chemical Engineering*, **18**, 713.
12. Savkovic Stevanovic J. (1996) Neural networks controller by inverse modeling for a distillation plant, *Computers&Chemical Engineering*, **20**, 925.
13. An P. E., M. Brown, Harris C. J. (1995) A global gradient noise covariance expression for stationary real Gaussian inputs, *IEEE Trans. Neural Networks*, **6**, 1549-1551.
14. Jordan M. I., Jacobs R. A. (1994) Hierarchical mixtures of experts and the EM algorithm, *Neural computation*, **6**, 181-214.
15. Savkovic Stevanovic J. (2019) Inference in stochastic information processing, *International Journal of Mathematical and Computational Methods*, **11**, 43-48.
16. Savkovic Stevanovic J. (2019) Neural-fuzzy artificial intelligence system, *International Journal of Control Systems and Robotics-IARAS*, **4**, 21-26.