# Probabilistic Modeling

In subject area: Computer Science

Probabilistic Modeling is a technique used in Computer Science to predict the likelihood of a given test sample belonging to a particular subject. It calculates the probability based on training samples and uses a graph model to find the probabilistic relationship between different nodes.

AI generated definition based on: Computers & Security, 2016

On this page

## Chapters and Articles

You might find these chapters and articles relevant to this topic.

Review article

## A survey on touch dynamics authentication in mobile devices

2016, Computers & Security

Pin Shen Teh, ... Ke Chen

### 6.1 Probabilistic modeling (PM)

The main idea behind the probabilistic modeling technique is to predict the likelihood of a given test sample belonging to a particular subject using the prior probability calculated from training samples (touch dynamics data acquired during user enrollment phase). One widely used probabilistic modeling technique is the Bayesian Network (Feng et al., 2013; Saravanan et al., 2014). It uses an acyclic graph model to find the probabilistic used as the parent node and the associated subject identity as a child node. Then, given a test sample (touch dynamics data acquired during user authentication phase), the intended child node is determined by the probability of the parent node (Jeanjaitrong and Bhattarakosol, 2013). Other variants of the probabilistic modeling technique include the Naive Bayes

View article

Read full article

URL: https://www.sciencedirect.com/science/article/pii/S0167404816300256

Chapter

# Meta-learning basics and background

2023, Meta-Learning

Lan Zou

## 1.7 Probabilistic modeling

Probabilistic modeling, also known as statistical modeling, is a statistical technique that accounts for random events in order to estimate potential outcomes of mathematical expressions across multiple random variables. The probabilistic method was proposed by the prolific Hungarian mathematician Paul Erdős in the late 1940s. This model relied on random variables together with the corresponding cumulative distribution functions to estimate a possible outcome regarding an event or a phenomenon by outputting a probability measure as a solution. Variables can be randomly sampled from standard probability distributions, including normal distribution (i.e., Gaussian distribution), uniform distribution, binomial distribution, Poisson distribution, and Bernoulli distribution. The Central Limited Theorem plays an essential role in formulating large-sized independent random variables into normal distributions.

Probabilistic modeling differs from machine learning and deep learning, as deep learning is an unknown unknown, machine learning is known known, but probabilistic modeling is known unknown. Foong et al. (2020) allowed enhanced dependencies in predictive distributions through a maximum-likelihood meta-learner called Convolutional Neural Process. Wu, Choi, Goodman, For further information on probabilistic modeling or statistic methods for machine learning, the following textbook is

View chapter

Explore book

Read full chapter

URL: https://www.sciencedirect.com/science/article/pii/B9780323899314000109

# Part II. More advanced machine learning schemes

2017, Data Mining (Fourth Edition)

Ian H. Witten, ... Christopher J. Pal

Chapter 9, Probabilistic methods, covers probabilistic modeling approaches that go far beyond the simple Naïve Bayes classifier introduced in Chapter 4, Algorithms: the basic methods. We begin with a review of some fundamental concepts, such as maximum likelihood estimation, that form the basis of probabilistic approaches. Then we examine Bayesian networks, a powerful way of extending the Naïve Bayes method to make it less "naïve" by accommodating datasets that have internal dependencies. Next we consider how clustering can be viewed from a probabilistic perspective as fitting a mixture of probability distributions to a dataset. This is essentially a form of density estimation, and we also discuss the alternative approach of kernel density estimation to model the distribution of a dataset.

Read full chapter

View PDF

Explore book

Read full chapter

URL: https://www.sciencedirect.com/science/article/pii/B9780128042915000210

Chapter

# Probabilistic methods

2017, Data Mining (Fourth Edition)

## 9.9 Further Reading and Bibliographic Notes

The field of probabilistic machine learning and data mining is enormous: it essentially subsumes all classical and modern statistical techniques. This chapter has focused on foundational concepts and some widely used probabilistic techniques in data mining and machine learning. Excellent books that focus on statistical and probabilistic methods include Hastie, Tibshirani, and Friedman (2009), and Murphy (2012). Koller and Friedman (2009)'s excellent book specializes in advanced techniques and principles of probabilistic graphical models.

The *K*2 algorithm for learning Bayesian networks was introduced by Cooper and Herskovits (1992). Bayesian scoring metrics are covered by Heckerman et al. (1995). Friedman, Geiger, and Goldszmidt (1997) introduced the tree augmented Naïve Bayes algorithm, and also describe multinets. Grossman and Domingos (2004) show how to use the conditional likelihood for scoring networks. Guo and Greiner (2004) present an extensive comparison of scoring metrics for Bayesian network classifiers. Bouckaert (1995) describes averaging over subnetworks. AODEs are described by Webb, Boughton, and Wang (2005), and AnDEs by Webb et al. (2012). AD trees were introduced and analyzed by Moore and Lee (1998)—the same Andrew Moore whose work on *k*D-trees and ball trees was mentioned in Section 4.10. Komarek

Show more

View chapter

Explore book

Read full chapter

URL: https://www.sciencedirect.com/science/article/pii/B978012804291500009X

Chapter

# Introduction to Medical Image Recognition, Segmentation, and Parsing

2016, Medical Image Recognition, Segmentation and Parsing

S.Kevin Zhou

View chapter

Explore book

Read full chapter

URL: https://www.sciencedirect.com/science/article/pii/B9780128025819000019

Review article

# A review on the computational approaches for gene regulatory network construction

2014, Computers in Biology and Medicine

Lian En Chai, ... Zalmiyah Zakaria

## 5.2 Applications of probabilistic Boolean network

Shmulevich et al. [44] constructed a genetic regulatory network model in which the four model classes were (i) rule-based dependencies between genes, (ii) a systematic study of global network dynamics, (iii) ability to handle uncertainty in data and model selection, and (iv) quantification of the relative influence and sensitivity of gene interactions. In the probabilistic context of Markov chain, the dynamics of the network were studied. The basic building blocks of Bayesian network is derived from probabilistic Boolean network that showed the probabilistic dependencies between genes and their parent genes.

Pal et al. [46] stated that automatic control can be employed to find optimal approaches to manipulating external variables. Dynamic evolution over a finite time horizon would be affected by the transition probabilities of an instantaneously random system. The extension of approaches of external control to context-sensitive probabilistic Boolean network is focused. Probabilistic Boolean network yield a subclass of Markovian corresponding to a probabilistic Boolean network is context-specific. It is composed of the current gene vector occupied as

Show more

View article

Read full article

URL: https://www.sciencedirect.com/science/article/pii/S0010482514000420

Review article

## • Big data acquisition and analysis • Development and differentiation

2018, Current Opinion in Systems Biology

Dominic Grün

### Towards a probabilistic understanding of cell fate emergence

Most available computational methods for the inference of lineage trees from single-cell RNA-seq data are deterministic in their assignment of each cell to an individual branch. This view is agnostic to the probabilistic nature of cell fate decision, assuming that a given progenitor state could give rise to a number of fates with different probabilities in a stochastic manner. Potentially, gene expression variability of master regulators could be an underlying mechanism, requiring that a random fluctuation of transcript levels crosses a given threshold in order to drive differentiation towards a particular fate [2,59–61]. A probabilistic modeling of cell differentiation in general leads to a better understanding of the commitment process by revealing the stages at which a progenitor looses potency for alternative fates. A beautiful example of this approach was implemented in the GPfates algorithm [62] for modeling the bifurcation into $T_H1$ and $T_{FH}$ sub-types of T helper cells during blood-stage *Plasmodium* infection in mice. After dimensional reduction and pseudotime inference within the Gaussian process framework, GPfates models cell states along a trajectory branching into multiple fates by a Gaussian Process Latent Variable Model, i.e. an overlapping mixture of Gaussian processes each corresponding to a distinct

The STEMNET algorithm [63] represents another approach to the probabilistic analysis of lineage priming. This supervised method relies on prior knowledge of terminal cell states, which can be, for example, unambiguously identified based on specific marker gene expression. STEMNET predicts the fate probability of naïve multipotent cells from these mature states by a robust elastic-net regularized general linear model. Applied to human hematopoietic cells, this algorithm predicts direct emergence of unilineage-restricted cells from low-primed hematopoietic stem and progenitor cells. Following a related strategy, the semi-supervised FateID algorithm [64] also starts from defined end states to learn the fate bias, i.e. the probability to differentiate into

View article

Read full article

URL: https://www.sciencedirect.com/science/article/pii/S2452310018300131

Review article

# Flood inundation modelling: A review of methods, recent advances and uncertainty analysis

2017, Environmental Modelling & Software

J. Teng, ... S. Kim

## 4.3 Simplified conceptual models

The main development in simplified conceptual models has been to address their deficiency of lacking mass conservation and representation of velocity. RFSM Dynamic uses continuity to distribute flood volume between storage areas and then computes the flow rates between these using Manning's equation. RFSM EDA employs a similar approach but is based on a diffusive approximation to shallow water equations. Attempts have been made to limit the volume on the floodplain using weightings derived from flow accumulation areas, as in the CSIRO TVD model (coincidently having the same name as the shock capturing scheme, but they represent entirely different modelling approaches). The enhancements of HAND model reportedly include enabling non-uniform inundation within a catchment, accounting for backwater effects in catchments with precipitation

As with the other model types, model assimilation using remote sensing data is another research hotspot that can improve the predictions from simplified conceptual models, largely by integration of remotely sensed flood extent and water depth during a model simulation. These methods can also benefit from

View article

Read full article

URL: https://www.sciencedirect.com/science/article/pii/S1364815216310040

Review article

## Machine learning applications in the resilience of interdependent critical infrastructure systems—A systematic literature review

2024, International Journal of Critical Infrastructure Protection

Basem A. Alkhaleel

### 4.2.2 During disruption applications

Among the early applications of ICISs resilience during the disruption phase is the work of [79] who proposed a probabilistic modeling scheme to analyze malevolent events that could appear in ICISs. The proposed scheme is based on modeling the relationship between datastreams coming from two network nodes using an USL approach known as the hidden Markov model (HMM). HMM is appropriate for dealing with dynamic systems as it can break the problem into a specific number of states which are connected in a probabilistic way. The goal of the model is to decide the state of the ICIs system under attack. [78] developed a framework that integrates various data-driven methods and physics-based approaches to help understand the propagation of failure between ICIs. In this framework, text mining approaches, where unstructured text is transformed into a structured format to identify meaningful patterns of information, are used to identify infrastructure failure patterns from gray data, including newspapers, media reports, Internet news, text data from social media, images, and videos. Using such data could compensate for the lack of official data on cascading failures. Note that the current to improve the preparedness of the system or current data to

Show more

View article

Read full article

URL: https://www.sciencedirect.com/science/article/pii/S1874548223000598

Review article

## Guest Editorial: Special issue on flexible and resilient urban energy systems

2023, International Journal of Electrical Power & Energy Systems

Tao Jiang, ... Vladimir Terzija

### 2.6 Recovery and restoration strategy of urban energy systems

As large amounts of small-scale distributed generation can provide emergency power supply to critical loads during blackouts, a dynamic microgrids-based load restoration model is proposed in [17] for resilient operation of urban power distribution systems during outages considering uncertain RESs power outputs and loads as well as droop-controlled DERs. In addition, an offline probabilistic modeling and online updating method is proposed to characterize the time-varying uncertainty of multiple RESs power outputs and loads. The case study shows that the proposed uncertainty modeling and updating method as well as the dynamic microgrids can increase the overall system resilience.

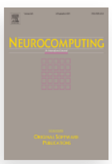
To further ensure the sustainability of energy continuity to urban critical loads, the authors in [18] propose a resiliency improving mechanism for distribution systems, which not only detects location and boundary of load outages, but also meets the energy constraints and data sharing limitations to maximize resiliency. In

View article

## Related terms:

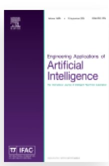
Decision-Making, Machine Learning, Bayesian Networks, Context-Free Grammars, Driving Behavior, Behavioral Approach, Clustering Algorithm, Probabilistic Approach, Bayesian Belief Network, Prediction Accuracy.

## Recommended publications



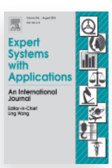

**Neurocomputing**

Journal



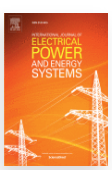

**Engineering Applications of Artificial Intelligence**

Journal




**Expert Systems with Applications**

Journal




**International Journal of Electrical Power & Energy Systems**

Journal

ELSEVIER