ISSN: 2321-7782 (Online)
Volume 2, Issue 11, November 2014
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

A Study on Hidden Markov Model (HMM)

Sumalatha.V[1]
Research Scholar
Bharathiar University
Asst. Professor, Vels University, India

Dr.Santhi.R[2]
Dean
Alpha Group of institutions
Chennai – India

Abstract: The Hidden Markov Model (HMM) was introduced by Crouse et al. in 1998 for modeling nonindependent, non Gaussian wavelet transform coefficient. They developed the equivalent of the forward-backward algorithm for hidden markov model. HMM is a widespread statistical model used in science, engineering and many other areas such as pattern recognition, speech recognition, machine translation, bioinformatics, computer vision, finance and economics and in social science. In this paper the contribution of forward-backward algorithm, Viterbi algorithm is applied with HMM. Furthermore examples of markov chain, markov tree, HMM are discussed.

Keywords: HMM, Viterbi algorithm, Forward – backward, Markov tree

I. Introduction

Markov Model:

In probability theory, a markov model is a stochastic model that assumes the markov property. A stochastic model models a process where the state depends on previous state in a non-deterministic way. A stochastic process has the markov property if the conditional probability distribution of future states of the process depends only upon the present state; that is, given the present, the future does not depend on the past. Markov model and their relationships are:

	System state is fully Observable	**System state is partially Observable**
System is Autonomous	Markov Chain	HMM
System is Controlled	Markov Decision Process	Partially observable Markov Decision process

Markov chain

The simplest markov model is the markov chain. It models the state of a system with a random variable that changes through a time. The distribution of variable depends on the previous state.

Model definition:

N states, $\{S_1, S_2, S_3, \ldots\ldots\ldots\ldots S_N\}$

Sequence of states $Q = \{q_1, q_2, q_3, \ldots\ldots q_N\}$

Initial probabilities $\pi = \{\pi_1, \pi_2, \pi_3, \ldots\ldots\ldots \pi_N\}$

$\Pi_i = P(q_1 = S_i)$

Transition matrix A N X N

$a_{ij} = P(q_{t+1} = S_j \mid q_t = S_i)$

Example of markov chain:

Weather Mode: 3 states

{ rainy,cloudy,sunny}

Problem: Forecast weather state based on the current weather state.

Markov Decision Process

A markov decision process is a markov chain in which state transition depends on the current state and an action vector that is applied to the system. It is used to compute a policy of action that will maximize some utility with respect to expected rewards. It is related to reinforcement learning and can be solved with value iteration and methods.

Partially observable markov decision process

A partially observable markov decision process is a markov chain process in which the state of the system is only partially observed. It is known to be NP complete. The example is controlling simple robots.

II. Hidden Markov Model (HMM)

A HMM is one in which to observe a sequence of emissions, but do not know the sequence of states the model went through to generate the emissions. Analyses of HMM seek to recover the sequence of states from the observed data. A Hidden Markov model (HMM) is a statistical Markov model in which the unobserved states are identified through Markov Process. HMM can be considered the simplest dynamic bayesian network. it is closely related to an early work of optimal nonlinear filtering problem called stochastic processes.

In simpler markov models like markov chain, the state is directly visible to the observer, and therefore the state transition probabilities are the parameters. In a Hidden Markov model, the state is not directly visible, but output, dependent on the state, is visible. Each state has a probability distribution over the possible output tokens. Therefore the sequence of tokens generated by an HMM gives some information about the sequence of states. Note that the adjective 'hidden' refers to the state sequence through which the model passes, not to the parameters of the model; the model is still referred to as a 'hidden' Markov model even if these parameters are known exactly.

Hidden markov model are especially known for their applications in temporal pattern recognition such as speech, handwriting, musical score following, partial discharges and bio informatics.

A hidden Markov model can be considered a generalization of a mixture model where the hidden variables which control the mixture component to be selected for each observation, are related through a Markov process rather than independent of each other.

To learn a HMM you have to specify its size (i.e. How many states and how many different observations are there). The simplest way is to let K-mean generate the initial hmm and then use baum-welch to optimal the initial HMM.

To make prediction using the learned HMM, you have to first find the state with the highest probability of being the one the system in right now. This is done using the Viterbi algorithm which identifies the most likely state sequence to have generated the given sequence of observations. Given this state you can calculate the most likely next observations.

Example: Coin Flipping

Assume that there are coins, one being biased towards heads (60% of the time it lands heads), one biased towards tails (60% of the time it lands tails). Assume further that a person picks up one of the coins and starts flipping it repeatedly (without you knowing which coin it is). From the sequence of results (heads or tails), learn a model that predicts if the next observation will be heads or tails. To do this (actually without knowing that the coins are biased or how biased they are) we can set up a

HMM learner that learns a HMM that models the process of flipping the coin. In this case we decide to let it learn a HMM with 2 states and 2 observations (heads =0, tails = 1). Two states should be sufficient there are only two possible coins.

States of Hidden Markov model

The prediction process is achieved by following states, Observation state, adjustment state and prediction state. The experiment results are encouraging and serve to show the promise of HMM in PSAP and they show accuracy in the next action prediction reaching up to 92%.

The HMM is a variant of a finite state machine having a set of hidden states, Q, an output alphabet (observations) , O, transition probabilities m A, output(emission) probabilities , B, and a initial state probabilities, Π. The current state is not observable. Instead each state produces an output with a certain probability (B). Usually the states, q, and outputs, O, are understood, so an HMM is asid to be triple, (A, B, Π)

Hidden states $\boldsymbol{Q} = \{ q_i \}, i = 1, . . . , N$.

Transition probabilities $\boldsymbol{A} = \{a_{ij} = P(q_j \text{ at } t+1 \mid q_i \text{ at } t)\}$, where $P(a \mid b)$ is the conditional probability of a given b, $t = 1, . . . , T$ is time, and q_i in $\boldsymbol{Q}$. Informally, $\boldsymbol{A}$ is the probability that the next state is q_j given that the current state is q_i. Observations (symbols) $\boldsymbol{O} = \{ o_k\}, k = 1, . . . , M$.

Emission probabilities $\boldsymbol{B} = \{ b_{ik}= b_i(o_k) = P(o_k \mid q_i) \}$, where $o_k$in $\boldsymbol{O}$. Informally, $\boldsymbol{B}$ is the probability that the output is $o_k$given that the current state is q_i.

Initial state probabilities $\boldsymbol{\Pi} = \{p_i= P(q_i \text{ at } t = 1)\}$.

III. Forward – Backward Algorithm

Using Forward & Backward algorithm for hidden markov model which computes the posterior marginals of all hidden state variables given a sequence of observations/emissions O1 :t = O1,---Ot for all hidden state variables Xk € {X1,....Xt}, the distribution P(Xk|O1:t).

This inference task is usually called smoothing. The algorithm makes use of the principle of dynamic programming to compute.

Efficiently the values that is required to obtain the posterior marginal distributions in two passes.

The first pass goes forward in time while the second goes backward in time; hence the forward-backward algorithm will performed.

The Forward-Backward Method

The forward method computes:

$$p(K_1, \ldots, K_{t-1}) = \sum_{i=1}^{N} \alpha_i(t)$$

The backward method computes (t>1):

.

$$p(K_t, \ldots, K_T) = \sum_{i=1}^{N} \beta_i(t)$$

We can do the forward-backward method which computes $p(K_1,\ldots,K_T)$ using formula (using any choice of $t=1,\ldots,T+1$!):

$$L = p(K_1, \ldots, K_T) = \sum_{I=1}^{N} \alpha_i(t)\beta_i(t)$$

IV. Viterbi Algorithm

Using viterbi algorithm we can find the path in a sequence of observed state. The Viterbi algorithm is a dynamic programming algorithm for finding the most likely sequence of hidden states – called the Viterbi path – that results in a sequence of observed events, especially in the context of Markov information sources and hidden Markov models.

Suppose we are given a Hidden Markov Model (HMM) with state space S, initial probabilities πi of being in state i and transition probabilities $a_{i,j}$ of transitioning from state i to state j. Say we observe outputs $Y_1. \dots Y_T$. The most likely state sequence $X_1. \dots X_T$ that produces the observations is given by the recurrence relations:

$V_{1,k} =$ P(y1 | k).π_k

$V_{t,k} =$ P(yt | k). $max_{x \in S}(a_{x,k} \cdot V_{t-1,x})$

Here $V_{t,k}$ is the probability of the most probable state sequence responsible for the first t observations that has k as its final state.

The Viterbi path can be retrieved by saving back pointers that remember which state x was used in the second equation. Let Ptr(k,t)be the function that returns the value of x used to compute $V_{t,k}$

if $t > 1$, or k if $t = 1$. Then:

$x_T = argmax_{x \in S}(V_{T,x})$

x_{t-1}= Ptr(x_t,t)

The complexity of this algorithm is

$O(T\ X\ |S|^2)$

V. Research Work Done in HMM

In Hidden markov model the famous author Lawrence R.Rabiner, IEEE had done a research in "A tutorial on HMM and selected applications in speech Recognition" in the year 1989.

Similarly "Generalized Hidden Markov model –theoritical framework" was done by Magdi A.Mohammed, Paul Gader Senior members of IEEE.

"Incremental learning of statistical motion pattern with growing HMM "by Dizan Vasquez , theirry Fraichard and Christian Laigier.

"Multiple observation HMM learning by aggregating ensemble models" by Nazanin Asadi, Abdolreza Mirzaei and Ehsan Haghshenas and

Behavior detection using confidence intervals of hidden Markov model by Richard R. Brooks, Jason M. Christopher Griffin, members of IEEE.

VI. Conclusion

In this paper "A study on hidden markov model" aims to know how hidden markov model is widely used through forward – backward algorithm, viterbi algorithmm to find the prediction methods using the observed states. Furthermore HMM is used to find the statistical analyses of death and birth ratio, prediction of diseases, genetical problems can be avoided through this methods.

References

1. Luger, George; Stubblefield, William (2004). Artificial Intelligence: Structures and Strategies for Complex Problem Solving (5th ed.). The Benjamin/Cummings Publishing Company, Inc. ISBN 0-8053-4780-1.
2. Neapolitan, Richard; Jiang, Xia (2012). Contemporary Artificial Intelligence. Chapman & Hall/CRC. ISBN 978-1-4398-4469-4.
3. Nilsson, Nils (1998). Artificial Intelligence: A New Synthesis. Morgan Kaufmann Publishers. ISBN 978-1-55860-467-4.
4. Russell, Stuart J.; Norvig, Peter (2003), Artificial Intelligence: A Modern Approach (2nd ed.), Upper Saddle River, New Jersey: Prentice Hall, ISBN 0-13-790395-2
5. Poole, David; Mackworth, Alan; Goebel, Randy (1998). Computational Intelligence: A Logical Approach. New York: Oxford University Press. ISBN 0-19-510270-3.
6. Winston, Patrick Henry (1984). Artificial Intelligence. Reading, Massachusetts: Addison-Wesley. ISBN 0-201-08259-4.
7. Rich, Elaine (1983). Artificial Intelligence. McGraw-Hill. ISBN 0-07-052261-8.
8. Rabiner, L. R., A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, Proceedings of the IEEE, vol. 77, no. 2, Feb. 1989, pgs 257 - 285. There is a lot of notation but verbose explanations accompany.
9. B..Pardo and W. Birmingham. Modeling Form for On-line Following of Musical Performances. AAAI-05 Proc., July 2005.
10. Satish L, Gururaj BI (April 2003). "Use of hidden Markov models for partial discharge pattern classification". IEEE Transactions on Dielectrics and Electrical Insulation.
11. Pr. Pieczynski, W. Pieczynski, Multisensor triplet Markov chains and theory of evidence, International Journal of Approximate Reasoning, Vol. 45, No. 1, pp. 1-16, 2007.
12. Boudaren et al., M. Y. Boudaren, E. Monfrini, W. Pieczynski, and A. Aissani, Dempster-Shafer fusion of multisensor signals in nonstationaryMarkovian context, EURASIP Journal on Advances in Signal Processing, No. 134, 2012.