Selective Prediction of Financial Trends with Hidden Markov Models

Yifu Huang

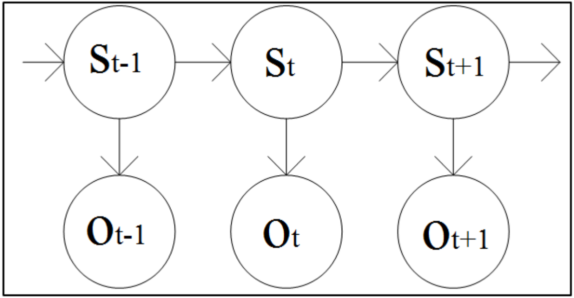
School of Computer Science, Fudan University
huangyifu@fudan.edu.cn

COMP630049 Statistical Learning and Dimensionality Reduction Report, 2014

Outline

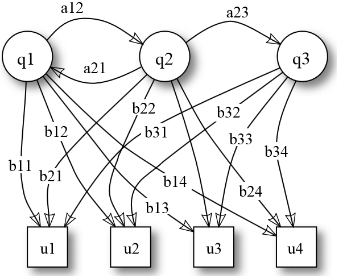
Hidden Markov Models - Concept

- A generative probabilistic model with latent states, where state transitions and observation emissions are assumed to be Markov processes
- Given an observation sequence $O = \{o_1, o_2, ..., o_T\}$ (each $o_t \in U$) that is generated by a HMM λ, we associate O with a latent state sequence $S = \{s_1, s_2, ..., s_T\}$ (each $s_t \in Q$) that most likely produces O



Hidden Markov Models - Definition

- HMM λ can be formally defined as a quintuple $\{N, M, \boldsymbol{\pi}, A, B\}$
  - N is the number of states; $Q=\{q_1, q_2, ..., q_N\}$
  - M is the number of observations; $U=\{u_1, u_2, ..., u_M\}$
  - $\boldsymbol{\pi}$ is the initial probability vector of states; $\pi_i=P(s_1 = q_i)$
  - A is the transition probability matrix of states
    - $a_{ij}=P(s_{t+1} = q_j | s_t = q_i)$
  - B is the observation emission probability matrix of states
    - $b_{ij}=P(o_t = u_j | s_t = q_i)$



Hidden Markov Models - Problem & Solution

- Problem 1: probability computation (compute $P(O|\lambda)$)
  - Solution: Forward Backward algorithm
- Problem 2: state annotation (maximize $P(S|\lambda, O)$)
  - Solution: Viterbi algorithm
- Problem 3: model training (maximize $P(O|\lambda)$)
  - Solution: Baum Welch algorithm

Selective Prediction - Definition

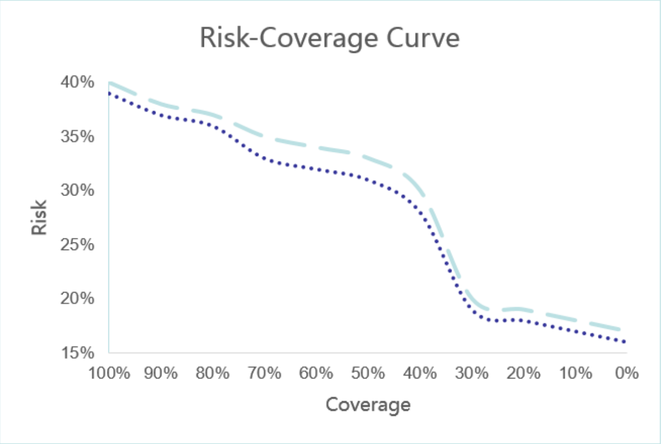
- Not ignorance, but ignorance of ignorance is the death of knowledge

$$Y_{t+1} = \begin{cases} F(X_t), & if\ G(X_t) = 1 \\ reject, & if\ G(X_t) = 0 \end{cases}$$

- A prediction framework that can qualify its own prediction results and reject the outputs when they are not confident enough

Selective Prediction - Evaluation

- Coverage $C = \frac{notREJECT}{ALL}$
- Risk $R = \frac{ERROR}{notREJECT}$
- Risk-Coverage Curve



Selective Hidden Markov Models - Definition - p_i

- Add state label p_i to each state q_i

Definition

Given an observation sequence $O=\{o_1, o_2, ..., o_T\}$ (indicating historical financial trend), a relative label sequence $L=\{l_1, l_2, ..., l_T\}$ (indicating next-day financial trend) and a HMM λ, the state label p_i denotes the most probable label that state q_i should have. Formally,

$$p_i = \arg \max_{l=up,down} \sum_{t=1, l_t=l}^{T} \gamma_{ti}. \tag{1}$$

Above, $\gamma_{ti} = P(s_t = q_i | O, \lambda)$ denotes the probability that the HMM λ stays at state q_i at time t, which can be efficiently computed by the forward-backward procedure.

Selective Hidden Markov Models - Definition - v_i

- Add empirical visit rate v_i to each state q_i

Definition

Given an observation sequence $O=\{o_1, o_2, ..., o_T\}$ (indicating historical financial trend) and a HMM λ, the empirical visit rate v_i denotes the fraction of time that the HMM λ spends at state q_i. Formally,

$$v_i = \frac{1}{T}\sum_{t=1}^{T} \gamma_{ti}. \tag{2}$$

Selective Hidden Markov Models - Definition - r_i

- Add empirical state risk r_i to each state q_i

Definition

Given an observation sequence $O=\{o_1, o_2, ..., o_T\}$ (indicating historical financial trend), a relative label sequence $L=\{l_1, l_2, ..., l_T\}$ (indicating next-day financial trend) and a HMM λ, the empirical state risk r_i denotes the rate of erroneous visits to state q_i. Formally,

$$r_i = \frac{\frac{1}{T}\sum_{t=1, l_t \neq p_i}^{T} \gamma_{ti}}{v_i}. \tag{3}$$

Selective Hidden Markov Models - Definition - *RS*

- Furthermore, we sort all HMM states by their empirical state risks in descending order and record them as $Q_d = \{q_{d_1}, q_{d_2}, ..., q_{d_N}\}$ (for each $j < k$, $r_{d_j} \geq r_{d_k}$).
- The low-quality HMM states, also called *reject states*, constitute the *reject subset RS*. Predictions at those states are prevented.

Definition

Given a coverage bound C_B, we label the reject states sequentially until their cumulative empirical visit rate $\sum_{j=1}^{K} v_{d_j}$ exceeds 1-C_B. Formally, the reject subset RS is defined as

$$RS = \{q_{d_1}, q_{d_2}, ..., q_{d_K} | \sum_{j=1}^{K} v_{d_j} \leq 1 - C_B, \sum_{j=1}^{K+1} v_{d_j} > 1 - C_B\} \tag{4}$$

Selective Hidden Markov Models - Definition - q_h

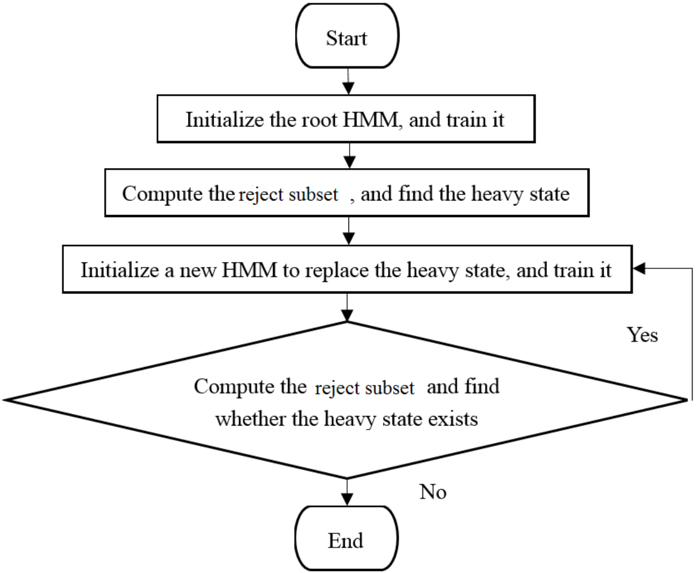
- Identify heavy state q_h

Definition

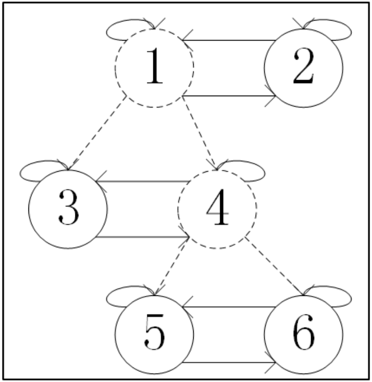
Given a visit bound V_B, state $q_{d_{K+1}}$ is identified as a heavy state q_h if its visit rate $v_{d_{K+1}} > V_B$.

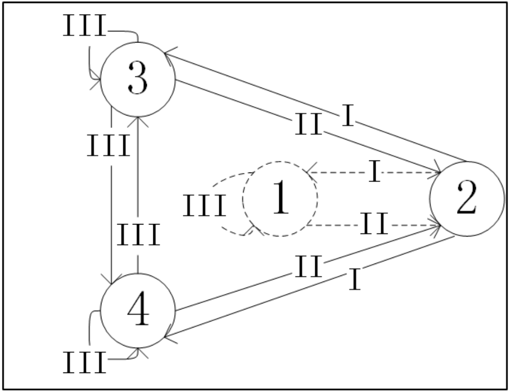
- The heavy state q_h is the cause of coarseness problem, and it should be recursively refined in the training stage.

Selective Hidden Markov Models - Training

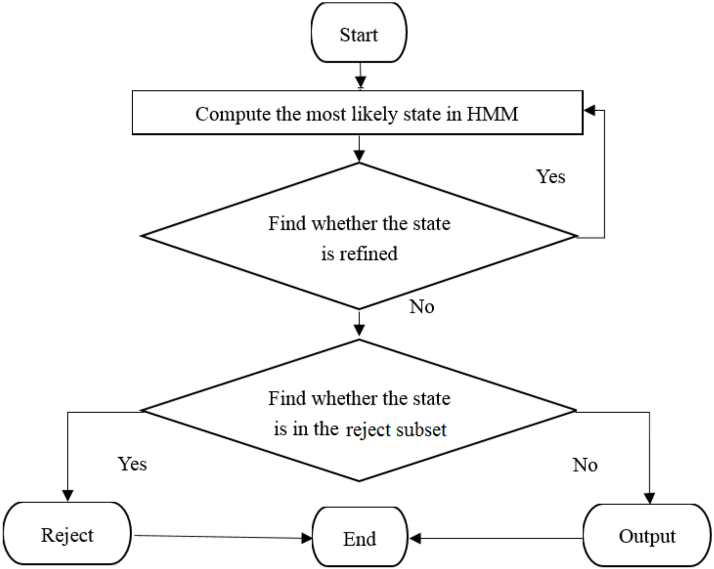


Selective Hidden Markov Models - Training - Refine





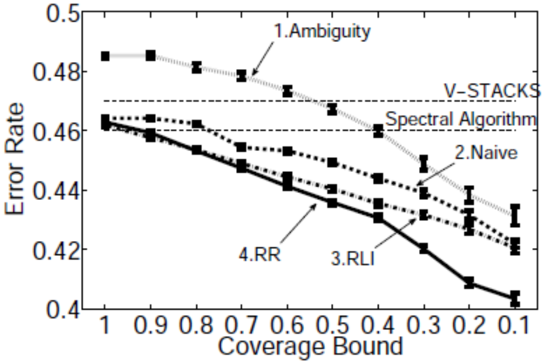
Selective Hidden Markov Models - Prediction



Experiment - Setup

- The data sequence in this experiment consisted of the 3000 S&P500 returns from 1/27/1999 to 12/31/2010
- Employe a walk-forward scheme in which the model is trained over the window of past W_p returns and then tested on the subsequent window of W_f "future" returns
  - $W_p = 2000$
  - $W_f = 50$
- Train and test using 30-fold crossvalidation, with each fold consisting of 10 random restarts

Experiment - Risk-Coverage Curve



(a) Error rate vs coverage bound

Discussion - Related Work

- Besides historical financial data, more and more additional indictors have been used to improve financial trend prediction
  - News reports [2][3]
  - Twitter mood [4][5]
  - Trading relationship [6]

The End

Thank you!

References I

- [1] Dmitry Pidan, Ran El-Yaniv: Selective Prediction of Financial Trends with Hidden Markov Models. NIPS 2011:855-863
- [2] Robert P. Schumaker, Hsinchun Chen: A Discrete Stock Price Prediction Engine Based on Financial News. IEEE Computer (COMPUTER) 43(1):51-56 (2010)
- [3] Xiaodong Li, Chao Wang, Jiawei Dong, Feng Wang, Xiaotie Deng, Shanfeng Zhu: Improving Stock Market Prediction by Integrating Both Market News and Stock Prices. DEXA 2011:279-293
- [4] Johan Bollen, Huina Mao, Xiao-Jun Zeng: Twitter mood predicts the stock market. J. Comput. Science (JOCS) 2(1):1-8 (2011)
- [5] Jianfeng Si, Arjun Mukherjee, Bing Liu, Qing Li, Huayi Li, Xiaotie Deng: Exploiting Topic based Twitter Sentiment for Stock Prediction. ACL 2013:24-29

References II

- [6] Xiao-Qian Sun, Hua-Wei Shen, and Xue-Qi Cheng. Trading network predicts stock price. Scientific Reports, 4(3711):1–6 (2014)