

HMM之Baum-Welch算法

原创

2016年11月10日 13:24:49

标签：机器学习 / 算法 / NLP / 统计学

876

注：本文中所有公式和思路来自于邹博先生的《机器学习升级版》，我只是为了加深记忆和理解写的本文。

前面介绍了HMM模型结构和HMM基本问题中的概率计算问题，本文介绍HMM基本问题中的参数学习问题。

如果训练数据包括观测序列和状态序列，则HMM的学习非常简单，是监督学习，如果只有观测序列的话，那么HMM的学习是需要使用EM算法的，是非监督学习。

监督学习：

根据大数定理（频率的极限是概率），我们可以轻易的得出下边的这几个结论：

初始概率：

$$\hat{\pi}_i = \frac{|q_i|}{\sum_i |q_i|}$$

转移概率：

$$\hat{a}_{ij} = \frac{|q_{ij}|}{\sum_{j=1}^N |q_{ij}|}$$

观测概率：

$$\hat{b}_{ik} = \frac{|s_{ik}|}{\sum_{k=1}^M |s_{ik}|}$$

上边给出的上个结论可以直接从给定的数据中数出来，没有任何难度。

非监督学习：

我们回顾一下EM算法的框架吧：



大数据学习路线



联系我们



请扫描二维码联系
webmaster@csdn.net
400-660-0111
QQ客服

关于 招聘 广告服务
©1999-2018 CSDN版权所有
京ICP证09002463号

经营性网站备案信息
网络110报警服务
中国互联网举报中心
北京互联网违法和不良信息举报中心

他的最新文章

- 压力测试工具siege的用法
- 【卷积神经网络-进化史】从LeNet
- linux 环境变量设置（临时 + 永久）
- gtest install
- linux(ubuntu)下的caffe编译安装

文章分类

- 机器学习
- linux
- deep learning

文章存档

- 2017年1月
- 2016年12月
- 2016年11月

他的热门文章

- linux(ubuntu)下的caffe编译安装 1748
- EM算法---基于隐变量的参数估计 1674
- 概率图模型之贝叶斯网络 1657

Repeat until convergence {

(E-step) For each i , set

$$Q_i(z^{(i)}) := p(z^{(i)}|x^{(i)}; \theta).$$

(M-step) Set

$$\theta := \arg \max_{\theta} \sum_{i=1}^n Q_i(z^{(i)}) \log \frac{p(x^{(i)}, z^{(i)}; \theta)}{Q_i(z^{(i)})}$$

}

所有的观测数据写成： $O = (o_1, o_2 \dots o_T)$ ，所有隐数据写成： $I = (i_1, i_2 \dots i_T)$ ，完全数据： $(O, I) = (o_1, o_2 \dots o_T, i_1, i_2 \dots i_T)$ ，完全数据的对数似然函数是： $\ln P(O, I | \lambda)$

假设 $\lambda \sim$ 是HMM参数的当前估计值， λ 是待估计的参数：

$$\begin{aligned} Q(\lambda, \bar{\lambda}) &= \sum_I (\ln P(O, I | \lambda)) P(I | O, \bar{\lambda}) \\ &= \sum_I \ln P(O, I | \lambda) \frac{P(O, I | \bar{\lambda})}{P(O, \bar{\lambda})} \\ &\propto \sum_I \ln P(O, I | \lambda) P(O, I | \bar{\lambda}) \end{aligned}$$

我们回顾一下EM的过程：首先需要预先给定的 $\lambda \sim$ 参数，然后带入 $P(I | O, \lambda \sim)$ ，然后将 $P(I | O, \lambda \sim)$ 带入 $Q(\lambda, \lambda \sim)$ 中，对对数似然函数 $\ln P(O, I | \lambda)$ 求期望，找到能使似然函数期望最大的新的 $\lambda \sim$ ，并将新的 $\lambda \sim$ 再次带回 $P(I | O, \lambda \sim)$ ，不断重复这个过程。

我们在前一篇文章中提到过暴力求解，并得到了最终的求解公式：

$$\begin{aligned} P(O, I | \lambda) &= P(O | I, \lambda) P(I | \lambda) \\ &= \pi_{i_1} b_{i_1 o_1} a_{i_1 i_2} b_{i_2 o_2} \cdots a_{i_{T-1} i_T} b_{i_T o_T} \end{aligned}$$

Q()函数可以写成：

$$\begin{aligned} Q(\lambda, \bar{\lambda}) &= \sum_I \ln P(O, I | \lambda) P(O, I | \bar{\lambda}) \\ &= \sum_I \ln \pi_{i_1} P(O, I | \bar{\lambda}) \\ &\quad + \sum_I \left(\sum_{t=1}^{T-1} \ln a_{i_t i_{t+1}} \right) P(O, I | \bar{\lambda}) \\ &\quad + \sum_I \left(\sum_{t=1}^T \ln b_{i_t o_t} \right) P(O, I | \bar{\lambda}) \end{aligned}$$

既然是要极大化Q，求得参数 π 、A、B

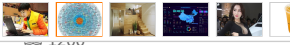
由于这三个参数分别在这三个项中，那么我们就可以分别极大化。

极大化初始概率：

加入CSDN，享受更精准的内容推荐，与500万程序员共同成长！



大数据学习路线



联系我们



请扫描二维码联系
webmaster@
400-660-011
QQ客服

关于 招聘 广告服务
©1999-2018 CSDN版权所有
京ICP证09002463号

经营性网站备案信息
网络110报警服务
中国互联网举报中心
北京互联网违法和不良信息举报中心

因为 π_i 满足加和为1，那么我们就可以利用拉格朗日乘法：

$$\sum_{i=1}^N \ln \pi_i P(O, i_1 = i | \bar{\lambda}) + \gamma \left(\sum_{i=1}^N \pi_i - 1 \right)$$

接着对 π 求偏导：

$$P(O, i_1 = i | \bar{\lambda}) + \gamma \pi_i = 0$$

接着我们可以对 i 求和：

$$\gamma = -P(O | \bar{\lambda})$$

最后将 γ 代回去得：

$$\pi_i = \frac{P(O, i_1 = i | \bar{\lambda})}{P(O | \bar{\lambda})} = \frac{P(O, i_1 = i | \bar{\lambda})}{\sum_{i=1}^N P(O, i_1 = i | \bar{\lambda})} = \frac{\gamma_1(i)}{\sum_{i=1}^N \gamma_1(i)}$$

注意：这个 γ 不是超参数的 γ ，而是我们之前求过的这么个小东西：

$$\gamma_i(i) = P(i_t = q_i | O, \lambda)$$

极大化转移概率和观测概率：

$$\sum_I \left(\sum_{t=1}^{T-1} \ln a_{i_t i_{t+1}} \right) P(O, I | \bar{\lambda}) = \sum_{i=1}^N \sum_{j=1}^N \sum_{t=1}^{T-1} \ln a_{ij} P(O, i_t = i, i_{t+1} = j | \bar{\lambda})$$

使用拉格朗日乘法：

$$a_{ij} = \frac{\sum_{t=1}^{T-1} P(O, i_t = i, i_{t+1} = j | \bar{\lambda})}{\sum_{t=1}^{T-1} P(O, i_t = i | \bar{\lambda})} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)}$$

同理可得：

$$b_{ik} = \frac{\sum_{t=1}^T P(O, i_t = i | \bar{\lambda}) I(o_t = v_k)}{\sum_{t=1}^T P(O, i_t = i | \bar{\lambda})} = \frac{\sum_{t=1, o_t=v_k}^T \gamma_t(i)}{\sum_{t=1}^T \gamma_t(i)}$$

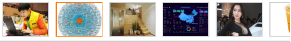
这几个结果就是可以直接写代码的，学习问题解决了，就剩最后一个预测问题了，下一篇文章就会介绍Viterbi算法。

版权声明：本文为博主原创文章，未经博主允许不得转载。 <https://blog.csdn.net/u012771351/article/details/53114829>

目前您尚未登录，请 [登录](#) 或 [注册](#) 后参与评论



大数据学习路线



联系我们



请扫描二维码联系
webmaster@csdn.net
400-660-0111
QQ客服

关于 招聘 广告服务
©1999-2018 CSDN版权所有
京ICP证09002463号

经营性网站备案信息
网络110报警服务
中国互联网举报中心
北京互联网违法和不良信息举报中心

前言在上一篇博文中，我简单地介绍了隐马尔科夫模型HMM，并且重点介绍了HMM的三个问题中的第一个，即概率计算问题。首先回顾一下这三个问题都是什么以及解决每个问题的主流算法： 概率计算问题即模型评价问题...

 njustzj001 2016年03月23日 21:34  11723

“ 相亲记 ”之从EM算法到Baum-Welch算法

“你还在期待爱情，但是先得解决婚姻呀 ”。



0







 firparks 2017年02月09日 19:39  1958





HMM之Baum-Welch算法

注：本文中所有公式和思路来自于邹博先生的《机器学习升级版》，我只是为了加深记忆和理解写的本文。 前面介绍过了HMM模型结构和HMM基本问题中的概率计算问题，本文介绍HMM基本问题中的参数学习...

 u012771351 2016年11月10日 13:24  876



Baum-Welch算法

Baum-Welch算法 精度下溢， 可以用标定或log来解决

 sadfasdgaaaasdfa 2015年07月12日 16:28  2841


隐马尔可夫模型之Baum-Welch算法详解

隐马尔可夫模型之Baum-Welch算法详解 在上篇博文中，我们学习了隐马尔可夫模型的概率计算问题和预测问题，但正当要准备理解学习问题时，发现学习问题中需要EM算法的相关知识，因此，上一周转而学习了...

 u014688145 2016年11月05日 18:03  14188



一秒创造无法计算的价值

每满2000返200，最高返5000元代金券





隐马尔可夫模型（ HMM ） 攻略

隐马尔可夫模型 (Hidden Markov Model，HMM) 最初由 L. E. Baum 和其它一些学者发表在一系列的统计学论文中，随后在语言识别，自然语言处理以及生物信息等领域体现了很大的价...

 likelet 2011年12月09日 10:44  172010



Baumwelch 算法训练HMM

学习问题 在HMM模型中，已知隐藏状态的集合S，观察值的集合O，以及一个观察序列（ o1,o2,...,on ），求使得该观察序列出现的可能性最大的模型参数（包括初始状态概率矩阵 π ，状态转移矩阵A，发射...

 linjiebelfast 2014年05月05日 17:14  1279



隐马尔可夫模型（ HMM ） 的MATLAB实现——Baum-Welch算法

Baum-Welch算法用来对隐马尔可夫模型的参数进行学习，Baum-Welch算法是EM算法的一种特例，属于非监督学习算法，下边第一部分程序的迭代次数为1的算法代码，其中调用了计算中所需要的两个变量...

 sinat_36005594 2017年03月31日 16:57  3116


Baum-Welch算法

最近 由于科研的需要，重新 学习了一下HMM里的学习和预测算法（ 解码 ），里面的两个重要的算法，就是Baum-Welch和Viterbi算法。概率计算问题即模型评价问题——前向算法和后向算法学习问题即参...

 zhaomengszu 2017年01月15日 14:58  565

隐马尔可夫(HMM)、前/后向算法、Viterbi算法 再次总结

本总结是是个人为防止遗忘而作，不得转载和商用。 说明：此篇是作者对“隐马尔可夫模型” 的第二次总结，因此可以算作对上次总结的查漏补缺以及更进一步的理解，所以很多在第一次总结中已经整理过的...

 xuevinaxue001 2016年09月01日 10:00  9061

加入CSDN，享受更精准的内容推荐，与500万程序员共同成长！



大数据学习路线



联系我们



请扫描二维码联系

 webmaster@csdn.net

 400-660-0111

 QQ客服 95017

关于 招聘 广告服务 

©1999-2018 CSDN版权所有
京ICP证09002463号

经营性网站备案信息
网络110报警服务
中国互联网举报中心
北京互联网违法和不良信息举报中心

 u011300443 2015年07月05日 17:41  8637

EM算法 实例讲解


第一次接触EM算法，是在完成半隐马尔科夫算法大作业时。我先在网
上下载了两份Baum-Welch算法的代码，通过复制粘贴，修修补补，用java实现了HMM算法（应用是韦小宝掷两种骰子的问题）。然后，参考...

程序员不会英语怎么行？

老司机教你一个数学公式秒懂天下英语



HMM及其算法（前向，Viterbi，Baum-Welch）

2012年11月26日 17:15 567KB 



隐马尔可夫模型 (Hidden Markov Model，HMM)理解

隐马尔可夫模型 (Hidden Markov Model，HMM) 最初由 L. E. Baum 和其它一些学者发表在一系列的统计学论文中，随后在语音识别，自然语言处理以及生物信息等领域体现了很大的价...

 qianwenhong 2014年11月26日 10:49  3497

隐马尔可夫模型（HMM）的MATLAB实现——前向后向算法

由于研究需要，在网上找了不少关于隐马尔可夫模型的MATLAB程序，可能是没有耐下心去看，总之感觉看懂别人写的程序很费劲，所以就自己动手写了一下。主要的参考书目是李航的《统计学习方法》，并实现了书中部...

 sinat_36005594 2017年03月31日 16:29  2867

隐马尔可夫模型（HMM）的MATLAB实现——Viterbi算法

维特比算法实际是用动态规划求解隐马尔可夫模型解码问题，即用动态规划求概率最大路径（最优路径）。代码中有示例，来自李航《统计学习方法》function [Delta,Psi,P,I] = Viter...

 sinat_36005594 2017年03月31日 17:08  3328

使用python实现HMM

 u014365862 2015年12月27日 16:50  3898

一直想用隐马尔可夫模型做图像识别，但是python的scikit-learn组件包的hmm module已经不再支持了，需要安装hmmlearn的组件，不过hmmlearn的多项式hmm每次出来的结果都...

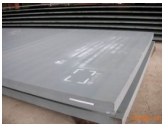
HMM-维特比算法原理和实现

 u012176591 2015年06月06日 15:01  1976

def viterbi(A,B,PI,O): T = np.shape(O)[0] N,K = np.shape(B) delta = np.zeros((T,N)) ...

耐磨板厂家

生产耐磨板的厂家



百度广告

白话机器学习算法（十六）HMM 隐马尔科夫链

隐马尔科夫链主要是两部分：

 wangxin110000 2014年04月04日 20:22  3204

HMM预测算法——Viterbi算法

 BD_Jiang 2016年09月18日 08:39  2134

HMM预测算法——Viterbi算法 HMM是一个生成模型，表示状态序列和观测序列的联合分布，但是状态序是隐藏的，不可观测的。Viterbi算法实际上用动态规划（dynamic pro...



大数据学习路线



联系我们



请扫描二维码联系
✉ webmaster@csdn.net
☎ 400-660-0111
🗣 QQ客服

关于 招聘 广告服务

©1999-2018 CSDN版权所有
京ICP证09002463号

经营性网站备案信息

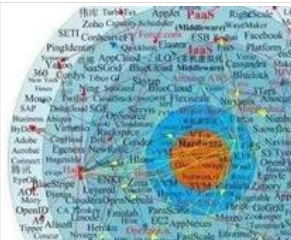
网络110报警服务

中国互联网举报中心

北京互联网违法和不良信息举报中心

理论知识 实现这个算法的时候遇到很多困难，因为当时连输入是什么，输出是什么都不知道。而且很多人把forward-backward算法和baum-welch两个算法等同起来了。所以我想先整理一下一些理论...

 allanmcgrady 2014年01月05日 16:25 3164



大数据学习路线



联系我们



请扫描二维码联系
✉ webmaster@csdn.net
☎ 400-660-0108
🗣 QQ客服 4006600108

关于 招聘 广告服务 网站地图
©1999-2018 CSDN版权所有
京ICP证09002463号

经营性网站备案信息
网络110报警服务
中国互联网举报中心
北京互联网违法和不良信息举报中心