Learning HMM Structure for Information Extraction

Speaker: Chunwei Yan

Begin

Introduction to HMM
Example

HMM 的使用

一个词一状态

Neighbor-merging 相邻合并

V-mergin

M-merging and Bayesia

买验结果

Conclude

Learning HMM Structure for Information Extraction

Speaker:

Chunwei Yan

互联网研发中心

March 8, 2013

Outline

Learning HMM Structure for Information Extraction

> Speaker: Chunwei Yan

Begin

Introduction to HMM Example

HMM 的使用

从训练集中学习 HMM 模型 一个词一状态

Neighbor-merging 相邻合

V-merging

1-merging and Bayesia

实验结果

- Begin
 - Introduction to HMM
 - Example
- ② HMM 的使用
 - 从训练集中学习 HMM 模型
 - 一个词一状态
 - Neighbor-merging 相邻合并
 - V-merging
 - M-merging and Bayesian
- ③ 实验结果
- Conclude

Learning HMM Structure for

Information

Extraction

Speaker:

Chunwei Y

Begin

Introduction to HMM

HMM 的使用

人训练集中学习 HMM 模型

一个词一状态

Neighbor-merging 相邻合法

V-mergii

M morging and Rawein

立验结里

Learning HMM Structure for Information Extraction

> Speaker: Chunwei Yar

Begin

Introduction to HMM

HMM 的使用

一个词一状态

Neighbor-merging 相邻合

V-mergin

M-merging and Bayesia

实验结果

Conclude

HMM 组成

- 状态 (states) 集合 Q {q_I, q₀, q₁, · · · q_F}
- 状态间的转移 (trasitions) $(q \rightarrow q')$
- 一个有限的观测 (output symbols) 集合 $\sum = \{\sigma_1, \sigma_2, \cdots, \sigma_M\}$

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yar

Begin

Introduction to HMM
Example

HMM 的使用

从训练集中学习 HMM 模型

Neighbor-merging 相邻台

M-merging and Bavesia

实验结果

. . .

HMM 组成

- 状态 (states) 集合 Q {q_I, q₀, q₁, · · · q_F}
- 状态间的转移 (trasitions) $(q \rightarrow q')$
- 一个有限的观测 (output symbols) 集合 $\sum = \{\sigma_1, \sigma_2, \cdots, \sigma_M\}$

模型系数

- 状态转移概率: P(q → q')
- 状态 q 观测为 σ 的概率: P(q↑σ)

Learning HMM Structure for Information Extraction

> Speaker: Chunwei Yan

Begin

Introduction to HMM

HMM 的使用

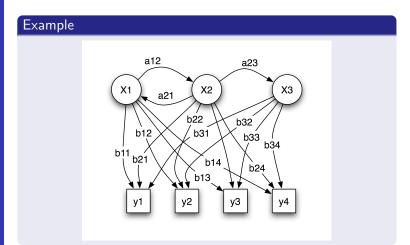
从训练集中学习 HMM 模 一个词一状态

Neighbor-merging 相邻合并

V-mergii

M-merging and Bayesian

实验结果



HMM 的使用

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yar

Begin

Introduction to HMM
Example

HMM 的使用

从训练集中学习 HMM 核

一个词一状态

V-----

M-merging and Bayesia

实验结果

Conclude

一个 HMM 模型观测为一个字符串 x 的概率

$$P(x|M) = \sum_{q_1, \dots, q_l \in Q^t} \prod_{k=1}^{l+1} P(q_{k-1} \to q_k) P(q_k \uparrow y_k) \quad (1)$$

HMM 的使用

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yar

Begin

Introduction to HMM Example

HMM 的使用

从训练集中学习 HMM 模型 一个词一状态

iveignbor-merging 4H %P &

V-mergii

M-merging and Bayesia

实验结果

Conclude

一个 HMM 模型观测为一个字符串 x 的概率

$$P(x|M) = \sum_{q_1, \dots, q_l \in Q^t} \prod_{k=1}^{l+1} P(q_{k-1} \to q_k) P(q_k \uparrow y_k)$$
 (1)

如下恢复出最可能生成观测 y 的状态序列

$$P(x|M) = \sum_{q_1, \dots, q_l \in Q^t} \prod_{k=1}^{l+1} P(q_{k-1} \to q_k) P(q_k \uparrow y_k) \quad (2)$$

Learning HMM Structure for Information Extraction

> Speaker: Chunwei Yan

Begin

Introduction to HMM

HMM 的使用

一个词一状态

Neighbor-merging 相邻合并

M marging and Rausci

实验结果

Conclude

论文头

Learning Hidden Markov Model Structure for Information Extraction

From: AAAI Technical Report WS-99-11. Compilation copyright © 1999, AAAI (www.aaai org), Ali fights reserved.

Kristie Seymore[©]
Kaymore[©]
Kaymore[©]
Ronald Rosenfeld[†]
Ronald Rosenfeld[†]
Ronald Rosenfeld[†]
Kaymore[©]
roni[©]
Rosenmu.edu

[†]School of Computer Science Carnegie Mellon University Pittsburgh, PA 15213 [‡]Just Research 4616 Henry Street Pittsburgh, PA 15213

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yar

Begir

Introduction to HMM

HMM 的使用

从训练集中学习 HMM 科

Neighbor-merging 相邻台

M-merging and Bayesi

实验结果

C = = = |...d=

论文头

Learning Hidden Markov Model Structure for Information Extraction

From: AAAI Technica Report WS-99-11. Compilation copyright © 1999, AAAI (www.aaai.org), Ali diphs reserved.

Kristie Seymore[†] Andrew McCallum^{††} Ronald Rosenfeld[†]
kseymore[©]ri.cmu.edu mccallum@justresearch.com coni@es.cmu.edu

[†]School of Computer Science Carnegie Mellon University Pittsburgh, PA 15213 [‡]Just Research 4616 Henry Street Pittsburgh, PA 15213

提取的 15 个类别

title, author, affiliation, address, note, email, date abstract, introduction, phone, keywords, web, degree, publication number(pubnum) and page

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yar

Begin

Introduction to HMM

HMM 的使用

从则称来中于刁 FIVIVI 保

1 141 17761

Neighbor-merging 相邻合

实验结果

Conclude

模型参数学习

可以的两种情况:

Learning HMM Structure for Information Extraction

> Speaker: Chunwei Yai

Begin

Introduction to HMM Example

HMM 的使用

从训练集中学习 HMM 模型

Neighbor-merging 相邻会

v-merging

实验结果

Conclude

模型参数学习

可以的两种情况:

● 每一个状态 (state) 代表一种类别 (class label), 比如标题, 作者, 地址等。

Learning HMM Structure for Information Extraction

> Speaker: Chunwai Va

Begin

Introduction to HMM Example

HMM 的使用

从训练集中字习 HMM 模型

Neighbor-merging 相邻合

M ----i-- --- D----i-

实验结果

Conclude

模型参数学习

可以的两种情况:

- 每一个状态 (state) 代表一种类别 (class label), 比如标题, 作者, 地址等。
- 一种类别联系到多种状态,每个状态间仅有有限的转移。

Learning HMM Structure for Information Extraction

> opeaker: Chunwei Ya

Begin

Introduction to HMM

HMM 的使用

一个词一状态

iveignbor-merging 4H

M-merging and Bavesia

实验结果

_ . .

模型参数学习

可以的两种情况:

- 每一个状态 (state) 代表一种类别 (class label), 比如标题, 作者, 地址等。
- 一种类别联系到多种状态,每个状态间仅有有限的转移。

标注新的论文头部

- 将论文头部的词作为观测值
- 用 Viterbi 算法恢复最有可能的状态序列

Learning HMM
Structure for
Information
Extraction

Speaker:

Begin

Introduction to HMM

HMM 的使用

从训练集中学习 HMM 模型

一个词一状态

Neighbor-merging 相邻合

V-mergin

M marging and Rayonia

实验结果

Conclude

首先要确定模型中有多少状态.

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yai

Begin

Introduction to HMM

HMM 的使用

从训练集中学习 HMM 模型

一个词一状态

V-mergii

M-merging and Bayesia

实验结果

Conclude

首先要确定模型中有多少状态.

可取方案:

Learning HMM
Structure for
Information
Extraction

Speaker: Shunwei Var

Begin

Introduction to HMM

....

I IIVIIVI LIJI文/T.

1 149 100.00

. .

M marging and Rayoria

从训练集中学习 HMM 模型

实验结果

Conclude

首先要确定模型中有多少状态.

可取方案:

• 仅仅为每一个状态分配一种类别

Learning HMM
Structure for
Information
Extraction

Chunwei Yan

Begin

Introduction to HMN

HMM 的使用

从训练集中学习 HMM 模型

M-merging and Bavesia

实验结果

Conclude

首先要确定模型中有多少状态.

可取方案:

- 仅仅为每一个状态分配一种类别
- 将一种类别联系到多个状态

统计方法

Learning HMM Structure for Information

Extraction

Speaker:

Begin

Introduction to HMM
Example

HMM 的使用

从训练集中学习 HMM 模型

. . . .

TACIBIDOS MICIBINE 1914

V-mergin

M-merging and Bayesia

实验结果

Conclude

为取得 HMM 系数的最大似然估计

$$\hat{P}(q \to q') = \frac{c(q \to q')}{\sum_{\sigma \in Q} c(q \to s)}$$
(3)

$$\hat{P}(q \uparrow \sigma) = \frac{c(q \uparrow \sigma)}{\sum_{\rho \in \Sigma} c(q \uparrow \rho)} \tag{4}$$

模型学习过程

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yan

Begin

Introduction to HMM
Example

HMM 的使用

从训练集中学习 HMM 模型

一个词一状态

V-merging

M-merging and Bayesian

实验结果

- 为每一个词独立分配一个状态
- ② neighbor-merging(相邻合并)
- ❸ 进一步合并
 - V-merging
 - M-merging
 - Bayesian model merging

为每一个词独立分配一个状态

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yar

Begin

Introduction to HMM
Example

HMM 的使用

11300候作市公司 1144

一个词一状态

Neighbor-merging 相邻合

V-mergi

M-merging and Bayesia

实验结果

Conclude

为训练集中每一个词分配一个独立的状态,同时,当 前词汇与下一个词汇间对应着状态转移

为每一个词独立分配一个状态

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yar

Begin

Introduction to HMM

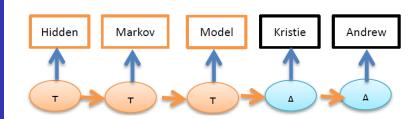
HMM 的使用

- 从训练集中学习 HMM 模型
- Neighbor-merging 相邻合
- V-merging
- M-merging and Bayesia

实验结果

Conclude

为训练集中每一个词分配一个独立的状态,同时,当 前词汇与下一个词汇间对应着状态转移



Neighbor-merging 相邻合并

Learning HMM Structure for Information Extraction

> Speaker: Chunwei Yar

Begin

Introduction to HMM
Example

HMM 的使用

从训练集中学习 HMM 模

Neighbor-merging 相邻合并

V morging

A-merging and Bavesia

实验结果

Conclude

合并共享转移 (transition) 以及相同类别 (class) 的所有状态

Neighbor-merging 相邻合并

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yar

Begin

Introduction to HMM

HMM 的使用

从训练集中学习 HMM 模型 一个词一状态

Neighbor-merging 相邻合并

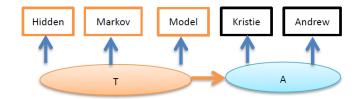
.

M-merging and Bayesian

实验结果

Conclude

合并共享转移 (transition) 以及相同类别 (class) 的所有状态



Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yan

Begin

Introduction to HMM Example

HMM 的使用

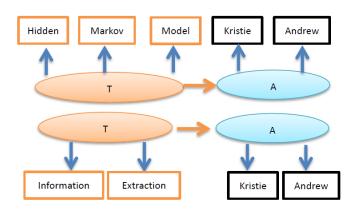
从训练集中学习 HMM 模型

Neighbor-merging 相邻合

V-mergi

M-merging and Bayesia

实验结果



Learning HMM
Structure for
Information
Extraction

- Speaker: Chunwei Yan

Begin

Introduction to HMM

HMM 的使用

人训练集中学习 HMM 模

Neighbor-marging 相邻会:

V-merging

M-merging and Bavesia

实验结果

Conclude

合并任何有相同类别标签且转移到相同状态或者从共 同状态转移的两个状态

Learning HMM Structure for Information Extraction

> Speaker: Chunwei Yar

Begin

Introduction to HMM

HMM 的使用

从训练集中学习 HMM 模型

Neighbor-merging 相邻合

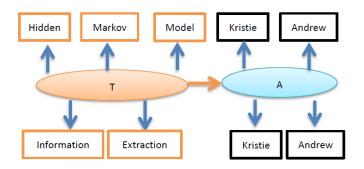
V-merging

M-merging and Bayesia

实验结果

Conclude

合并任何有相同类别标签且转移到相同状态或者从共 同状态转移的两个状态



Learning HMM Structure for Information Extraction

Speaker:

hunwei Yar

Begir

Introduction to HMM

HMM 的使用

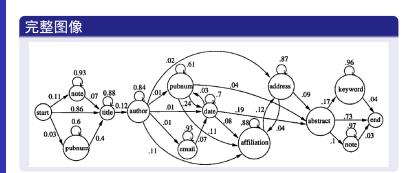
从训练集中学习 HMM 模型

Neighbor-merging 相邻台

V-merging

M-merging and Bayesian

实验结果



M-merging and Bayesian Model Merging

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yar

Begin

Introduction to HMM

HMM 的使用

从训练集中学习 HMM 模型

一个词一状态

Neighbor-merging 相邻合并

V-mergin

M-merging and Bayesian

实验结果

Conclude

多个状态对应一个类别:

M-merging and Bayesian Model Merging

Learning HMM Structure for Information Extraction

> Speaker: Chunwei Yan

Begin

Introduction to HMM

HMM 的使用

从训练集中学习 HMM 模

一个词一状态

V-merging

M-merging and Bayesian

实验结果

Conclude

多个状态对应一个类别:

M-merging

在 neighbor-merging之后,利用人工的方法迭代地合并状态.

M-merging and Bayesian Model Merging

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yar

Begin

Introduction to HMM

HMM 的使用

从训练集中学习 HMM 模型

Neighbor-merging 相邻合

V-merging

M-merging and Bayesian

实验结果

Conclude

多个状态对应一个类别:

M-merging

在 neighbor-merging之后,利用人工的方法迭代地合并状态.

Bayesian Model Merging

利用贝叶斯方法进行 HMM 的模型状态合并

其它方面

Learning HMM
Structure for
Information
Extraction

Speaker:

Begin

Introduction to HMM Example

HMM 的使用

从训练集中学习 HMM:

N - 11 - 1900

Meighbor-merging 12 49 E

M-merging and Bayesian

实验结果

Conclude

Unlabeled Data(未标注数据)

在 V-merging 或 M-merging 的基础之上,采用

Baum-Welch算法进行非监督学习.

其它方面

Learning HMM
Structure for
Information
Extraction

Speaker:

Begin

Introduction to HMM

HMM 的使用

从训练集中学习 HMM 模

Neighbor-merging 相邻合

V-mergir

M-merging and Bayesian

实验结果

C = = = |...d =

Unlabeled Data(未标注数据)

在 V-merging 或 M-merging 的基础之上, 采用 Baum-Welch 算法进行非监督学习.

Distantly-labeled Data(远亲标注数据)

为充实训练集,混合使用已经标注的相似数据集.

- L+D: Labeled 和 Dislabeled 数据集简单组合
- L*D: Labeled 和 Dislabeled 数据集的线性插值

实验数据

Learning HMM
Structure for
Information
Extraction

Speaker:

Regin

Introduction to HMM

HMM 的使日

从训练集中学习 HMM 模型

一个词一状态

V-merging

M-merging and Bayesia

实验结果

Type	Source	Word Tokens
Labeled	500 headers	23,557
Unlabeled	5,000 headers	287,770
Distantly-labeled	176 BibTeX files	2,390,637

一个类别一个状态的实验结果

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yai

Begin

Introduction to HMM

HMM 的使用

从训练集中学习 HMM 模型

Neighbor-merging 相邻合

·

M-merging and Bayesia

实验结果

			Accuracy		
Model	# states	# trans	L	L+D	L*D
full	17	255	62.8	57.4	64.5
self	17	252	85.9	83.1	89.4
ML	17	149	90.5	89.4	92.4
smooth	17	255	89.9	88.8	92.0

M-merging V-Merging 结果

Learning HMM
Structure for
Information
Extraction

Speaker: Shunwei Yar

Begin

Introduction to HMM

HMM 的使用

从训练集中学习 HMM 模型

Neighbor-merging 相邻合

v-merging

实验结果

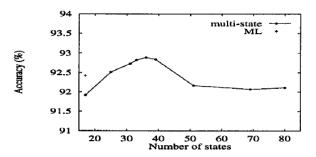


Figure 2: Extraction accuracy for multi-state models as states are merged.

M-merging V-Merging 结果

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yar

Begin

Introduction to HMM

HMM 的使用

从训练集中学习 HMM 模型 一个词一状态

Neighbor-merging 相邻合

实验结果

Conclude

			Accuracy		
Model	# states	# trans	L	L+D	L*D
ML	17	149	90.5	89.4	92.4
M-merged	36	164	91.3	90.5	92.9
V-merged	155	402	90.6	89.7	92.7

Table 3: Extraction accuracy (%) for models learned from data compared to the best model that uses one state per class.

Baum-Welch 后期加工结果

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yai

Begin

Introduction to HMM
Example

HMM 的使用

从训练集中学习 HMM 模型

Neighbor-merging 相邻合:

实验结果

Conclude

	ML		M-merged		
	Acc.	PP	Acc.	PP	
initial	92.4	471	92.9	482	
$\lambda = 0.5$	90.1	374	89.4	361	
λ varies	89.7	364	88.8	349	

Table 4: Extraction accuracy (%) and test set perplexity (PP) for the ML and M-merged models after Baum-Welch training.

Learning HMM
Structure for
Information

Extraction
Speaker:

. Chunwei Y

Begin

Introduction to HMM

Example

HMM 的使用

从训练集中学习 HMM 模型

一个词一状态

Neighbor-merging 相邻会主

V-mergii

立验结里

Learning HMM
Structure for
Information
Extraction

Speaker:

Begin

Introduction to HMM

HMM 的使用

从训练集中学习 HMM 模

一行词一从窓

Neighbor-merging 相邻旨

M-merging and Bayesia

实验结果

Conclude

● HMM 在信息提取领域真的有潜力, 这篇论文中达到了 92.9% 的精度.

Learning HMM Structure for Information Extraction

> Speaker: Chunwei Yar

Begin

Introduction to HMM

HMM 的使用

一个词一状态

Neighbor-merging 相邻合

M-merging and Bayesia

实验结果

- HMM 在信息提取领域真的有潜力, 这篇论文中达到了 92.9% 的精度.
- ② 一个类别 (class) 对应一到多个状态 (state) 更加有效

Learning HMM Structure for Information Extraction

Speaker:

Begir

Introduction to HMM
Example

HMM 的使用

- 从训练集中学习 HMM 模型 一个词一状态
- Neighbor-merging 相等旨
- M-merging and Bayesia

实验结果

- HMM 在信息提取领域真的有潜力, 这篇论文中达到了 92.9% 的精度.
- ② 一个类别 (class) 对应一到多个状态 (state) 更加有效
- ⑤ 远亲标注 (Distantly-labeled) 数据在模型系数估计方面有很大帮助, 插值组合

References

Learning HMM
Structure for
Information
Extraction

Speaker: Chunwei Yar

Begin

Introduction to HMM
Example

HMM 的使用

从训练集中学习 HMM 模型 一个词一状态

Neighbor-merging 相邻合并

V-merging

M-merging and Bavesia

实验结果

Conclude



[Kristie .S, Andrew .M and Ronald .R, 1999]

Learning Hidden Markov Model Structure for

Information Extraction

School of Computer Science Carneie Mellon University

Just Research