**基于HMM的命名实体识别**

## 学习任务

学习任务是命名实体识别。

在文本中对实体的引用可以有三种形式：命名性指称、名称性指称、代词性指称。例如在句子“中国乒乓球男队主教练刘国梁出席了会议，他布置了备战世乒赛的具体安排”中“刘国梁”的指称项有三个，其中“中国乒乓球男队主教练”是名词性指称，“刘国梁”是命名性指称，“他”是代词性指称。

命名实体识别任务是识别出文本中实体的命名性指称项，并标明其类别。一般来说，命名实体识别任务中，实体类别一般分为三大类（实体类、时间类、数字类）、七小类（人名、机构名、地名、时间、日期、货币和百分比）。不同任务对命名实体类别粒度的需求不同，例如在有些任务中只要识别出一个实体是人，有些任务则需要识别出一个实体是教师、学生或者医生。

## 学习算法

## 数据集

数据集是从网上获取的一个简历数据集，分为三个文件，分别是train、test、devlopment.

数据集中的实体类别有NAME(名字)、CONT()、EDU(学历)、TITLE()、ORG()、RACE()、PRO()、LOC()、

## 学习结果

## 理论分析