## 02-main

## 2023年5月21日

## 议员在议案上的投票记录预测其所属党派

为了研究不同党派的议员对议题选择的偏向性,某国国会记录了每个议员对每个议题的投票历史, 如图所示。现在想根据某个国会议员对每个议题的投票情况,预测该国会议员属于哪个党派。

- 1. 完成数据集读取
- 2. 完成数据集的缺失值填充,把所有放弃投票的值填充为字符串 None
- 3. 将所有议题作为分类特征
- 4. 构建分类独热编码器
- 5. 完成数据分类模型构建及训练
- 6. 对模型进行 10 折交叉检验
- 7. 完成数据分类预测

## 在 Github 中查看

1

```
[11]: import pandas as pd
[12]: data = pd.read_csv('议案投票.csv')
      data.head()
[12]:
                    Name Party Campaign Finance Overhaul
         ID
      0
          1
             Abercrombie
                             D
                                                        Y
                                                          \
                                                        Y
      1
          2
                Ackerman
      2
          3
                Aderholt
                             R
                                                        N
      3
          4
                    Akin
                             R
                                                        N
                Allen T.
        Unemployment and Tax Benefits Fiscal 2003 Budget Resolution
      0
                                    N
                                                                   N
                                                                      \
```

N

N

| 2  |  | Y                                      |  | Y               |  |
|--|--|--|--|-----------------|--|
| 3  |  | Y                                      |  | Y               |  |
| 4  |  | N                                      |  | N               |  |
|  |  |  |  |                 |  |
| Perman   | ent Tax Cuts Food                      | Stamps Nuclear                         | Waste  |                 |  |
| 0  | N                                      | Y                                      | N \  |                 |  |
| 1  | N                                      | Y                                      | N  |                 |  |
| 2  | Y                                      | N                                      | Y  |                 |  |
| 3  | Y                                      | N                                      | Y  |                 |  |
| 4  | N                                      | Y                                      | Y  |                 |  |
|  |  |  |  |                 |  |
| Fiscal   | 2003 Defense Aut                       | horization E                           | Estate Tax Repea                             | .1              |  |
| 0  |  | N                                      |  | Υ \             |  |
| 1  |  | N                                      |  | N               |  |
| 2  |  | Υ                                      |  | Y               |  |
| 3  |  | Υ                                      |  | Y               |  |
| 4  |  | N                                      |  | N               |  |
|  |  |  |  |                 |  |
|  |  |  |  |                 |  |
| Marrie   | ed Couples Tax Rel                     | ief Late Term A                        | Abortion Ban Hom                             | eland Sec/Unio  | n Memb   |
| Marrie<br>O  | ed Couples Tax Rel                     | ief Late Term A                        | Abortion Ban Hom<br>N                        | eland Sec/Unio  | n Memb<br>N \  |
|  | ed Couples Tax Rel                     |  |  | eland Sec/Unio  |  |
| 0  | d Couples Tax Rel                      | Y                                      | N  | eland Sec/Unio  | N \  |
| 0  | ed Couples Tax Rel                     | Y<br>N                                 | N<br>N                                       | eland Sec/Unio  | N \<br>N   |
| 0<br>1<br>2  | ed Couples Tax Rel                     | Y<br>N<br>Y                            | N<br>N<br>Y                                  | eland Sec/Unio  | N \<br>N<br>Y  |
| 0<br>1<br>2<br>3<br>4                                    |  | Y<br>N<br>Y<br>Y                       | N<br>N<br>Y<br>Y                             |                 | N \ N Y Y  |
| 0<br>1<br>2<br>3<br>4<br>Homela                          | ed Couples Tax Rel                     | Y N Y Y N N                            | N<br>N<br>Y<br>Y                             |                 | N \ N Y Y N  |
| 0<br>1<br>2<br>3<br>4<br>Homela                          |  | Y N Y Y N Tice Emp Homelar             | N<br>N<br>Y<br>Y                             |                 | N \ N Y Y N  |
| 0<br>1<br>2<br>3<br>4<br>Homela<br>0                     |  | Y N Y Y N rice Emp Homelar Y           | N<br>N<br>Y<br>Y                             |                 | N \ N Y Y N N  |
| 0<br>1<br>2<br>3<br>4<br>Homela<br>0<br>1                |  | Y N Y Y N Tice Emp Homelar Y Y         | N<br>N<br>Y<br>Y                             |                 | N \ N Y Y N N  The second of t |
| 0<br>1<br>2<br>3<br>4<br>Homela<br>0<br>1<br>2<br>3      |  | Y N Y Y N Tice Emp Homelar Y Y N       | N<br>N<br>Y<br>Y                             |                 | N \ N Y Y N N  The second of t |
| 0<br>1<br>2<br>3<br>4<br>Homela<br>0<br>1                |  | Y N Y Y N Tice Emp Homelar Y Y         | N<br>N<br>Y<br>Y                             |                 | N \ N Y Y N N  The second of t |
| 0<br>1<br>2<br>3<br>4<br>Homela<br>0<br>1<br>2<br>3<br>4 | and Sec/Civil Serv                     | Y N Y Y N rice Emp Homelar Y Y N N N N | N<br>Y<br>Y<br>N<br>nd Sec/Whistlebl         | ower Protectio  | N \ N Y Y N N  The second of t |
| 0 1 2 3 4 Homela 0 1 2 3 4 Andean                        | and Sec/Civil Serv                     | Y N Y Y N rice Emp Homelar Y Y N N N Y | N Y Y N ad Sec/Whistlebl                     | ower Protection | N \ N Y Y N N  The second of t |
| 0 1 2 3 4 Homela 0 1 2 3 4 Andean                        | and Sec/Civil Serv<br>Trade Abortion S | Y N Y Y N rice Emp Homelar Y Y N N N Y | N Y Y N nd Sec/Whistlebl                     | ower Protection | N \ N Y Y N N  The second of t |
| 0 1 2 3 4 Homela 0 1 2 3 4 Andean                        | and Sec/Civil Serv                     | Y N Y Y N rice Emp Homelar Y Y N N N Y | N Y Y N  and Sec/Whistlebl  s Medical Malpra | ower Protection | N \ N Y Y N N  The second of t |

```
3
                   Y
                                             N
                                                                         Y
      4
                   N
                                              Υ
                                                                         N
        Military Support for UN Resolution
      0
                                         N
      1
      2
                                         N
      3
                                         N
      4
                                         Y
      [5 rows x 23 columns]
[13]: features = data.columns[3:].tolist()
[14]: from sklearn.preprocessing import OneHotEncoder
[15]: one_hot_encoder = OneHotEncoder()
      one_hot_encoder.fit(data[features])
      one_hot_data = one_hot_encoder.transform(data[features])
[16]: from sklearn.naive_bayes import BernoulliNB
[17]: BNBModel = BernoulliNB()
[18]: from sklearn.model_selection import cross_val_score
[19]: cvs = cross_val_score(
          BNBModel,
          one_hot_data,
          data['Party'],
          cv=10
      cvs.mean()
```

[19]: 0.9908562367864693