

محمدرضا غفرانی  
۴۰۰۱۳۱۰۷۶  
۱۲ مرداد ۱۴۰۱

# مدل‌های گرافیکی احتمال

## پروژه درس

### نکات مهم در پیاده‌سازی پروژه

- در هنگام بارگذاری داده‌های مربوط به سری‌های زمانی تنها ستون‌های ویژگی (یعنی  $f_i$ ) به همراه ستون target خوانده می‌شود. در ادامه سطرهایی که مقدار ستون target آن‌ها برابر ۱ و ۲ نیست حذف می‌شود. همچنین یک واحد از تمامی اعدادی که در ستون target قرار دارند کم می‌شود تا مقادیر target برابر اعداد صفر و یک شود. در نهایت مقادیر خانه‌های خالی با استفاده از مد داده‌های ستون متناظر پر می‌شود.

- با توجه به آن که تعداد داده‌های کمی در دسترس است بنابراین در اغلب سناریوها فرض شده است که مقدار ویژگی‌ها از هم مستقل هستند. به عبارت دیگر داریم:

$$P(f_1, f_2, f_3, f_4) = \prod_{i=1}^4 P(f_i)$$

- با توجه به طول زیاد سری‌های زمانی، برای آن که محاسبات از لحاظ عددی پایدار باشد؛ بیشتر الگوریتم‌ها بر پایه لگاریتم پیاده‌سازی شده‌اند.

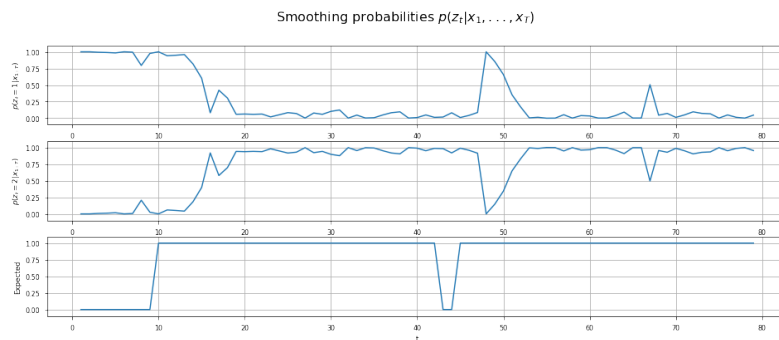
- یکی از چالش‌های پیاده‌سازی این پروژه مسئله متغیر بودن طول داده‌های سری‌های زمانی بود. در این پیاده‌سازی چون توابع مهمی نظیر `alpha_recursion`، `smoothing`، `beta_recursion` و `pair_marginals` در سطح یک سری زمانی عمل می‌کنند، بنابراین مشکلی از این بابت ندارند. مهم‌ترین مشکل از بابت الگوریتم Expectation Maximization است، که آن هم با توجه به اطلاعات موجود در این لینک می‌تواند از روی تک‌تک داده‌ها محاسبه شود و در نتیجه مشکلی از این بابت ندارد.

- لینک‌های زیر مراجع اصلی هستند که برای پیاده‌سازی الگوریتم‌های بیان شده استفاده شده‌اند.

- مرجع اول
- مرجع دوم
- مرجع سوم

## نمودار مقدار $\gamma$

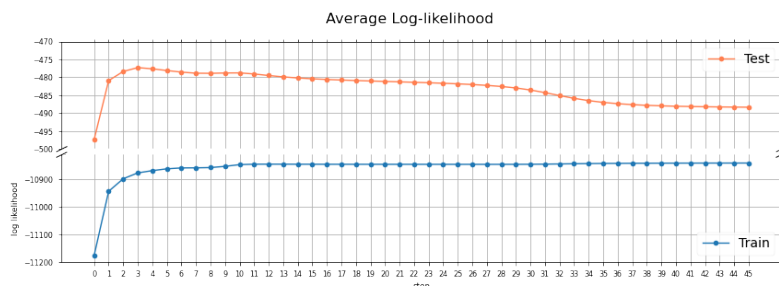
مقدار  $\gamma$  برای داده آزمون در شکل ۱ آورده شده است. نمودار سوم نیز برچسب داده‌ای است که نمودار برای آن رسم شده است. همان‌طور که مشاهده می‌شود نمودار  $\gamma$  ویژگی‌های اولیه‌ای نظیر  $P(z_1|x_{1:r}) + P(z_2|x_{1:r}) = 1$  را دارا بوده و پیش‌بینی می‌کند که برچسب داده از زمان صفر تا حدود ۲۰ برابر ۱ و از آن به بعد برابر ۲ است. البته در گام‌های زمانی نزدیک ۵۰ برچسب بعضی از گام‌ها ۱ برچسب زده شده است.



شکل ۱: نمودار مقادیر  $\gamma$

## نمودار Expectation Maximization

روند همگرایی likelihood به ازای داده‌های آموزشی و آزمون در شکل ۲ مشاهده می‌شود. آستانه توقف الگوریتم مرحله‌ای است که مجموع تغییرات هر یک از پارامترهای  $\pi$  و  $M$  به صورت مجزا از  $10^{-3}$  کوچک‌تر باشد. همان‌طور که مشاهده می‌شود  $\log$  likelihood روند صعودی را پیموده است در حالی که گرچه روی داده‌های تست روند کلی حرکت نیز صعودی است اما در نقاطی این حرکت به حرکت نزولی تبدیل شده است. این اتفاق نشان می‌دهد که الگوریتم روی داده‌های آموزشی بیش‌برازش شده است.



شکل ۲: نمودار همگرایی الگوریتم Expectation Maximization

## الگوریتم Viterbi

شبه کد الگوریتم Viterbi در شبه کد ۱ آورده شده است. شیوه عمل الگوریتم به صورت گرافیکی نیز در شکل ۳ مشاهده می‌شود. الگوریتم Viterbi برای تولید محتمل‌ترین برچسب برای یک دنباله زمانی در هر گام محتمل‌ترین حالت رسیدن به حالت فعلی از حالت قبلی را با استفاده از احتمال انتقال حالت‌ها به هم و احتمال تولید خروجی محاسبه کرده و ذخیره می‌کند. پس از یافتن محتمل‌ترین حالت به ازای تمامی گام‌های زمانی، کار برچسب‌گذاری با دنبال کردن محتمل‌ترین حالت قبلی انجام می‌شود.

---

### Algorithm 1 شبه کد الگوریتم Viterbi

---

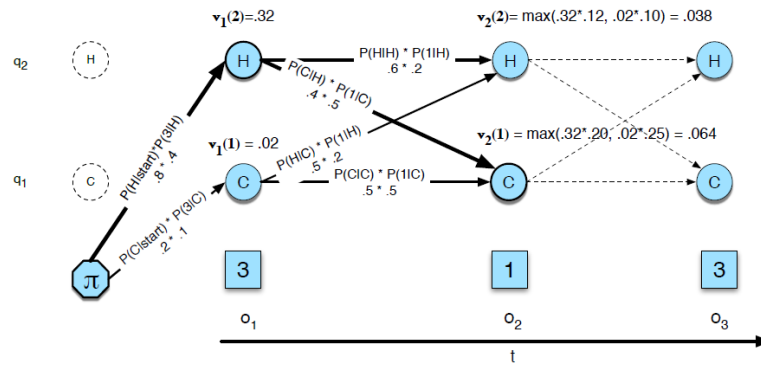
```

procedure Viterbi( $O, S, \Pi, Y, A, B$ )
  for each state  $i = 1, 2, \dots, K$  do
     $T_1[i, 1] \leftarrow \pi_i \cdot B_{iy_1}$ 
     $T_2[i, 1] \leftarrow 0$ 
  end for
  for each observation  $j = 2, 3, \dots, K$  do
    for each observation  $i = 1, 2, 3, \dots, K$  do
       $T_1[i, j] \leftarrow \max_k (T_1[k, j-1] \cdot A_{ki} \cdot B_{iy_j})$ 
       $T_2[i, j] \leftarrow \arg \max_k (T_1[k, j-1] \cdot A_{ki} \cdot B_{iy_j})$ 
    end for
  end for
   $z_T \leftarrow \arg \max_k (T_1[k, T])$ 
   $x_T \leftarrow s_{z_T}$ 
  for  $j = T, T-1, \dots, 2$  do
     $z_{j-1} \leftarrow T_2[z_j, j]$ 
     $x_{j-1} \leftarrow s_{z_{j-1}}$ 
  end for
  return  $X$ 
end procedure

```

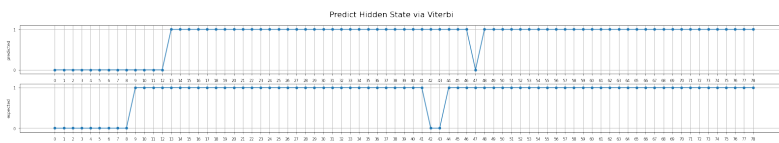
---

در شکل ۴ نمونه‌ای از اجرای الگوریتم viterbi بر روی یک دنباله زمانی آورده شده است. پارامترهای این الگوریتم با استفاده از شمارش حالت‌ها در داده‌های آموزشی به دست آمده است. همان‌طور که مشاهده می‌شود الگوریتم viterbi توانسته است بیشتر برچسب‌ها را به درستی پیش‌بینی کند. ابتدای دنباله را با توجه به احتمال‌های اولیه ( $\pi = [1, 0]$ ) به درستی برچسب یک پیش‌بینی کرده است. در ادامه چند گام زمانی پشت سر هم را به درستی برچسب یک پیش‌بینی کرده است. این اتفاق با توجه به این که مقدار درایه 0,0 در ماتریس  $M$  نسبت به درایه 0,1 بیشتر است بنابراین این رفتار نیز طبیعی است. از گام زمانی شماره ۱۲ با توجه به خروجی سری زمانی حالت سیستم از یک به ۲ تغییر کرده و در همان حالت باقی مانده است. این اتفاق به صورت مشابه با باقی ماندن در حالت یک توجیه می‌شود. در ادامه روند الگوریتم برچسب پیش‌بینی ۲ بوده است که دلیل این اتفاق بیشتر بودن مقدار 1,1



شکل ۳: شیوه عمل الگوریتم Viterbi

نسبت به 0, 1 در ماتریس  $M$  است. این اتفاق ادامه داشته و تنها در گام شماره ۴۷ با توجه به خروجی گام‌های زمانی الگوریتم حالت یک را پیش‌بینی کرده است.



شکل ۴: عملکرد الگوریتم viterbi

## قسمت ر

با توجه به تعداد زیاد متغیرهای زمینه‌ای و کوچک بودن مجموعه داده بررسی حضور تنها یکی از متغیرهای زمینه‌ای بررسی شده است. البته این بررسی صرفاً به حضور یک ویژگی زمینه‌ای محدود نشده و حالت‌های مختلف ترکیب ویژگی‌ها نیز بررسی شده است. این بررسی در جدول ۱ آورده شده است. با مقایسه نتایج این جدول با نتایج جدول ۲ مشخص می‌شود که متغیرهای زمینه‌ای عموماً اندکی عملکرد مدل پایه را بهبود می‌دهند. دلیل این اتفاق این است که چون ما مقادیر  $M$  و  $\epsilon$  را به ازای هر مقدار زمینه‌ای جداگانه محاسبه کرده‌ایم بنابراین این تخمین‌ها برای داده مد نظر دقیق‌تر بوده و در نتیجه به دقت بهتری می‌رسد. البته پیش‌بینی می‌شود که اگر متغیرهای زمینه‌ای بیشتری در مدل اعمال می‌شد عملکرد مدل افت می‌کرد. چرا که با افزایش متغیرهای زمینه‌ای احتمال بیش‌برازش مدل روی داده‌های آموزشی بیشتر شده و در نتیجه عملکرد مدل روی داده‌های آموزشی افت می‌کند.

LT

Table 1: بررسی تاثیر حضور متغیرهای زمینه‌ای

Context	Features	Accuracy
c1	(0,)	0.7770
c1	(1,)	0.3933
c1	(2,)	0.6212
c1	(3,)	0.6704
c1	(4,)	0.4806
c1	(0, 1)	0.7647
c1	(0, 2)	0.7645
c1	(0, 3)	0.7645
c1	(0, 4)	0.7781
c1	(1, 2)	0.5737
c1	(1, 3)	0.6441
c1	(1, 4)	0.5049
c1	(2, 3)	0.6545
c1	(2, 4)	0.6694
c1	(3, 4)	0.6847
c1	(0, 1, 2)	0.7816
c1	(0, 1, 3)	0.7785
c1	(0, 1, 4)	0.7709
c1	(0, 2, 3)	0.7714
c1	(0, 2, 4)	0.7826
c1	(0, 3, 4)	0.7802
c1	(1, 2, 3)	0.6326
c1	(1, 2, 4)	0.6889
c1	(1, 3, 4)	0.6779
c1	(2, 3, 4)	0.7157
c1	(0, 1, 2, 3)	0.7869
c1	(0, 1, 2, 4)	0.8056
c1	(0, 1, 3, 4)	0.7914
c1	(0, 2, 3, 4)	0.7771
c1	(1, 2, 3, 4)	0.6696
c1	(0, 1, 2, 3, 4)	0.7793
c2	(0,)	0.7268
c2	(1,)	0.4104
c2	(2,)	0.5400
c2	(3,)	0.6342
c2	(4,)	0.5401
c2	(0, 1)	0.7181
c2	(0, 2)	0.7263

c2	(0, 3)	0.7789
c2	(0, 4)	0.7292
c2	(1, 2)	0.5957
c2	(1, 3)	0.6383
c2	(1, 4)	0.5270
c2	(2, 3)	0.6140
c2	(2, 4)	0.6270
c2	(3, 4)	0.6460
c2	(0, 1, 2)	0.7514
c2	(0, 1, 3)	0.7647
c2	(0, 1, 4)	0.7324
c2	(0, 2, 3)	0.7473
c2	(0, 2, 4)	0.7214
c2	(0, 3, 4)	0.7499
c2	(1, 2, 3)	0.6233
c2	(1, 2, 4)	0.6379
c2	(1, 3, 4)	0.6577
c2	(2, 3, 4)	0.6568
c2	(0, 1, 2, 3)	0.7477
c2	(0, 1, 2, 4)	0.7326
c2	(0, 1, 3, 4)	0.7466
c2	(0, 2, 3, 4)	0.7666
c2	(1, 2, 3, 4)	0.6660
c2	(0, 1, 2, 3, 4)	0.7552
c3	(0,)	0.7574
c3	(1,)	0.5394
c3	(2,)	0.6383
c3	(3,)	0.6669
c3	(4,)	0.4809
c3	(0, 1)	0.7867
c3	(0, 2)	0.7613
c3	(0, 3)	0.7611
c3	(0, 4)	0.7704
c3	(1, 2)	0.6734
c3	(1, 3)	0.7279
c3	(1, 4)	0.6255
c3	(2, 3)	0.6573
c3	(2, 4)	0.6866
c3	(3, 4)	0.6865
c3	(0, 1, 2)	0.7886
c3	(0, 1, 3)	0.7821
c3	(0, 1, 4)	0.8104

c3	(0, 2, 3)	0.7679
c3	(0, 2, 4)	0.7796
c3	(0, 3, 4)	0.7820
c3	(1, 2, 3)	0.6848
c3	(1, 2, 4)	0.7371
c3	(1, 3, 4)	0.7688
c3	(2, 3, 4)	0.6923
c3	(0, 1, 2, 3)	0.7776
c3	(0, 1, 2, 4)	0.8083
c3	(0, 1, 3, 4)	0.8036
c3	(0, 2, 3, 4)	0.7890
c3	(1, 2, 3, 4)	0.7507
c3	(0, 1, 2, 3, 4)	0.8021
c4	(0,)	0.7399
c4	(1,)	0.4104
c4	(2,)	0.5750
c4	(3,)	0.6711
c4	(4,)	0.4789
c4	(0, 1)	0.7922
c4	(0, 2)	0.7416
c4	(0, 3)	0.7460
c4	(0, 4)	0.7638
c4	(1, 2)	0.6075
c4	(1, 3)	0.6152
c4	(1, 4)	0.5361
c4	(2, 3)	0.6400
c4	(2, 4)	0.6553
c4	(3, 4)	0.6912
c4	(0, 1, 2)	0.7539
c4	(0, 1, 3)	0.7842
c4	(0, 1, 4)	0.8106
c4	(0, 2, 3)	0.7469
c4	(0, 2, 4)	0.7478
c4	(0, 3, 4)	0.7685
c4	(1, 2, 3)	0.6209
c4	(1, 2, 4)	0.6540
c4	(1, 3, 4)	0.6617
c4	(2, 3, 4)	0.6589
c4	(0, 1, 2, 3)	0.7626
c4	(0, 1, 2, 4)	0.7633
c4	(0, 1, 3, 4)	0.7898
c4	(0, 2, 3, 4)	0.7500

c4	(1, 2, 3, 4)	0.6525
c4	(0, 1, 2, 3, 4)	0.7529
c5	(0,)	0.7601
c5	(1,)	0.4770
c5	(2,)	0.6417
c5	(3,)	0.6482
c5	(4,)	0.5303
c5	(0, 1)	0.7665
c5	(0, 2)	0.7678
c5	(0, 3)	0.7662
c5	(0, 4)	0.7641
c5	(1, 2)	0.5744
c5	(1, 3)	0.6089
c5	(1, 4)	0.5693
c5	(2, 3)	0.6284
c5	(2, 4)	0.6727
c5	(3, 4)	0.6738
c5	(0, 1, 2)	0.7695
c5	(0, 1, 3)	0.7715
c5	(0, 1, 4)	0.7776
c5	(0, 2, 3)	0.7750
c5	(0, 2, 4)	0.7756
c5	(0, 3, 4)	0.7811
c5	(1, 2, 3)	0.6221
c5	(1, 2, 4)	0.6560
c5	(1, 3, 4)	0.6784
c5	(2, 3, 4)	0.6929
c5	(0, 1, 2, 3)	0.7706
c5	(0, 1, 2, 4)	0.7705
c5	(0, 1, 3, 4)	0.7800
c5	(0, 2, 3, 4)	0.7864
c5	(1, 2, 3, 4)	0.6938
c5	(0, 1, 2, 3, 4)	0.7669
c6	(0,)	0.7633
c6	(1,)	0.3974
c6	(2,)	0.6329
c6	(3,)	0.6479
c6	(4,)	0.4761
c6	(0, 1)	0.7583
c6	(0, 2)	0.7712
c6	(0, 3)	0.7726
c6	(0, 4)	0.7689



c6	(1, 2)	0.6255
c6	(1, 3)	0.6762
c6	(1, 4)	0.5318
c6	(2, 3)	0.6353
c6	(2, 4)	0.6959
c6	(3, 4)	0.7168
c6	(0, 1, 2)	0.7717
c6	(0, 1, 3)	0.7887
c6	(0, 1, 4)	0.7648
c6	(0, 2, 3)	0.7844
c6	(0, 2, 4)	0.7791
c6	(0, 3, 4)	0.7812
c6	(1, 2, 3)	0.6604
c6	(1, 2, 4)	0.7008
c6	(1, 3, 4)	0.7171
c6	(2, 3, 4)	0.7170
c6	(0, 1, 2, 3)	0.7736
c6	(0, 1, 2, 4)	0.8018
c6	(0, 1, 3, 4)	0.7975
c6	(0, 2, 3, 4)	0.7802
c6	(1, 2, 3, 4)	0.7117
c6	(0, 1, 2, 3, 4)	0.7876
c7	(0,)	0.7344
c7	(1,)	0.5913
c7	(2,)	0.6255
c7	(3,)	0.6552
c7	(4,)	0.5314
c7	(0, 1)	0.7802
c7	(0, 2)	0.7367
c7	(0, 3)	0.7341
c7	(0, 4)	0.7554
c7	(1, 2)	0.6758
c7	(1, 3)	0.6867
c7	(1, 4)	0.6347
c7	(2, 3)	0.6236
c7	(2, 4)	0.7042
c7	(3, 4)	0.6926
c7	(0, 1, 2)	0.7948
c7	(0, 1, 3)	0.8140
c7	(0, 1, 4)	0.7939
c7	(0, 2, 3)	0.7538
c7	(0, 2, 4)	0.7510

c7	(0, 3, 4)	0.7566
c7	(1, 2, 3)	0.6921
c7	(1, 2, 4)	0.7062
c7	(1, 3, 4)	0.7060
c7	(2, 3, 4)	0.6877
c7	(0, 1, 2, 3)	0.8041
c7	(0, 1, 2, 4)	0.8014
c7	(0, 1, 3, 4)	0.8201
c7	(0, 2, 3, 4)	0.7625
c7	(1, 2, 3, 4)	0.7390
c7	(0, 1, 2, 3, 4)	0.8056
c8	(0,)	0.7646
c8	(1,)	0.4076
c8	(2,)	0.5905
c8	(3,)	0.6701
c8	(4,)	0.4841
c8	(0, 1)	0.7712
c8	(0, 2)	0.7668
c8	(0, 3)	0.7731
c8	(0, 4)	0.7794
c8	(1, 2)	0.6010
c8	(1, 3)	0.6257
c8	(1, 4)	0.5242
c8	(2, 3)	0.6268
c8	(2, 4)	0.6788
c8	(3, 4)	0.6685
c8	(0, 1, 2)	0.7691
c8	(0, 1, 3)	0.7719
c8	(0, 1, 4)	0.7872
c8	(0, 2, 3)	0.7681
c8	(0, 2, 4)	0.7836
c8	(0, 3, 4)	0.7732
c8	(1, 2, 3)	0.6373
c8	(1, 2, 4)	0.6789
c8	(1, 3, 4)	0.6901
c8	(2, 3, 4)	0.6917
c8	(0, 1, 2, 3)	0.7657
c8	(0, 1, 2, 4)	0.7856
c8	(0, 1, 3, 4)	0.7817
c8	(0, 2, 3, 4)	0.7809
c8	(1, 2, 3, 4)	0.7026
c8	(0, 1, 2, 3, 4)	0.7892

c9	(0,)	0.7393
c9	(1,)	0.4505
c9	(2,)	0.5640
c9	(3,)	0.6648
c9	(4,)	0.4818
c9	(0, 1)	0.7815
c9	(0, 2)	0.7392
c9	(0, 3)	0.7397
c9	(0, 4)	0.7610
c9	(1, 2)	0.5750
c9	(1, 3)	0.5991
c9	(1, 4)	0.5430
c9	(2, 3)	0.6244
c9	(2, 4)	0.6997
c9	(3, 4)	0.6966
c9	(0, 1, 2)	0.7824
c9	(0, 1, 3)	0.7804
c9	(0, 1, 4)	0.8009
c9	(0, 2, 3)	0.7491
c9	(0, 2, 4)	0.7598
c9	(0, 3, 4)	0.7650
c9	(1, 2, 3)	0.6016
c9	(1, 2, 4)	0.6794
c9	(1, 3, 4)	0.6795
c9	(2, 3, 4)	0.6907
c9	(0, 1, 2, 3)	0.7644
c9	(0, 1, 2, 4)	0.7948
c9	(0, 1, 3, 4)	0.7999
c9	(0, 2, 3, 4)	0.7603
c9	(1, 2, 3, 4)	0.6856
c9	(0, 1, 2, 3, 4)	0.7844
c10	(0,)	0.7501
c10	(1,)	0.4357
c10	(2,)	0.5959
c10	(3,)	0.6248
c10	(4,)	0.5031
c10	(0, 1)	0.7466
c10	(0, 2)	0.7559
c10	(0, 3)	0.7683
c10	(0, 4)	0.7843
c10	(1, 2)	0.5860
c10	(1, 3)	0.5788

c10	(1, 4)	0.5099
c10	(2, 3)	0.6218
c10	(2, 4)	0.6829
c10	(3, 4)	0.6741
c10	(0, 1, 2)	0.7469
c10	(0, 1, 3)	0.7742
c10	(0, 1, 4)	0.7840
c10	(0, 2, 3)	0.7752
c10	(0, 2, 4)	0.7915
c10	(0, 3, 4)	0.7920
c10	(1, 2, 3)	0.5823
c10	(1, 2, 4)	0.6494
c10	(1, 3, 4)	0.6497
c10	(2, 3, 4)	0.6750
c10	(0, 1, 2, 3)	0.7705
c10	(0, 1, 2, 4)	0.7861
c10	(0, 1, 3, 4)	0.7930
c10	(0, 2, 3, 4)	0.7974
c10	(1, 2, 3, 4)	0.6725
c10	(0, 1, 2, 3, 4)	0.7794
c10.1	(0,)	0.7497
c10.1	(1,)	0.5315
c10.1	(2,)	0.5951
c10.1	(3,)	0.6682
c10.1	(4,)	0.4830
c10.1	(0, 1)	0.7743
c10.1	(0, 2)	0.7523
c10.1	(0, 3)	0.7554
c10.1	(0, 4)	0.7567
c10.1	(1, 2)	0.5989
c10.1	(1, 3)	0.6400
c10.1	(1, 4)	0.6045
c10.1	(2, 3)	0.6316
c10.1	(2, 4)	0.6901
c10.1	(3, 4)	0.6983
c10.1	(0, 1, 2)	0.7798
c10.1	(0, 1, 3)	0.7838
c10.1	(0, 1, 4)	0.7516
c10.1	(0, 2, 3)	0.7587
c10.1	(0, 2, 4)	0.7617
c10.1	(0, 3, 4)	0.7644
c10.1	(1, 2, 3)	0.6292

c10.1	(1, 2, 4)	0.7070
c10.1	(1, 3, 4)	0.6900
c10.1	(2, 3, 4)	0.6978
c10.1	(0, 1, 2, 3)	0.7727
c10.1	(0, 1, 2, 4)	0.7870
c10.1	(0, 1, 3, 4)	0.7982
c10.1	(0, 2, 3, 4)	0.7685
c10.1	(1, 2, 3, 4)	0.7154
c10.1	(0, 1, 2, 3, 4)	0.7745
c12	(0,)	0.8011
c12	(1,)	0.4249
c12	(2,)	0.5990
c12	(3,)	0.6393
c12	(4,)	0.5278
c12	(0, 1)	0.7694
c12	(0, 2)	0.8019
c12	(0, 3)	0.7996
c12	(0, 4)	0.8271
c12	(1, 2)	0.5723
c12	(1, 3)	0.6059
c12	(1, 4)	0.5554
c12	(2, 3)	0.6328
c12	(2, 4)	0.7037
c12	(3, 4)	0.7114
c12	(0, 1, 2)	0.7640
c12	(0, 1, 3)	0.7730
c12	(0, 1, 4)	0.7874
c12	(0, 2, 3)	0.7889
c12	(0, 2, 4)	0.8204
c12	(0, 3, 4)	0.8222
c12	(1, 2, 3)	0.6120
c12	(1, 2, 4)	0.7136
c12	(1, 3, 4)	0.6852
c12	(2, 3, 4)	0.7039
c12	(0, 1, 2, 3)	0.7641
c12	(0, 1, 2, 4)	0.7867
c12	(0, 1, 3, 4)	0.7935
c12	(0, 2, 3, 4)	0.8109
c12	(1, 2, 3, 4)	0.6906
c12	(0, 1, 2, 3, 4)	0.7805
c13	(0,)	0.7581
c13	(1,)	0.5174

c13	(2,)	0.5928
c13	(3,)	0.6614
c13	(4,)	0.4784
c13	(0, 1)	0.7534
c13	(0, 2)	0.7542
c13	(0, 3)	0.7680
c13	(0, 4)	0.7676
c13	(1, 2)	0.5960
c13	(1, 3)	0.6417
c13	(1, 4)	0.5786
c13	(2, 3)	0.6234
c13	(2, 4)	0.6722
c13	(3, 4)	0.6841
c13	(0, 1, 2)	0.7775
c13	(0, 1, 3)	0.7982
c13	(0, 1, 4)	0.7874
c13	(0, 2, 3)	0.7689
c13	(0, 2, 4)	0.7600
c13	(0, 3, 4)	0.7700
c13	(1, 2, 3)	0.6463
c13	(1, 2, 4)	0.7137
c13	(1, 3, 4)	0.6938
c13	(2, 3, 4)	0.6870
c13	(0, 1, 2, 3)	0.7920
c13	(0, 1, 2, 4)	0.7802
c13	(0, 1, 3, 4)	0.7973
c13	(0, 2, 3, 4)	0.7742
c13	(1, 2, 3, 4)	0.7453
c13	(0, 1, 2, 3, 4)	0.7723
c14	(0,)	0.7454
c14	(1,)	0.3743
c14	(2,)	0.5954
c14	(3,)	0.6786
c14	(4,)	0.5231
c14	(0, 1)	0.7348
c14	(0, 2)	0.7501
c14	(0, 3)	0.7657
c14	(0, 4)	0.7586
c14	(1, 2)	0.6091
c14	(1, 3)	0.6536
c14	(1, 4)	0.5358
c14	(2, 3)	0.6429

c14	(2, 4)	0.6519
c14	(3, 4)	0.6698
c14	(0, 1, 2)	0.7471
c14	(0, 1, 3)	0.7444
c14	(0, 1, 4)	0.7437
c14	(0, 2, 3)	0.7605
c14	(0, 2, 4)	0.7687
c14	(0, 3, 4)	0.7674
c14	(1, 2, 3)	0.6202
c14	(1, 2, 4)	0.6384
c14	(1, 3, 4)	0.6757
c14	(2, 3, 4)	0.6797
c14	(0, 1, 2, 3)	0.7504
c14	(0, 1, 2, 4)	0.7384
c14	(0, 1, 3, 4)	0.7407
c14	(0, 2, 3, 4)	0.7627
c14	(1, 2, 3, 4)	0.6737
c14	(0, 1, 2, 3, 4)	0.7580

## قسمت ز

برای پاسخ به این قسمت تمامی ترکیب‌های مختلف بودن و نبودن ویژگی‌ها بررسی شد (نتایج این بررسی در جدول ۲ آمده است). و در نهایت به این نتیجه رسیدیم که برخلاف انتظار تاثیر دادن متغیر پنجم باعث بهتر شدن عملکرد الگوریتم viterbi می‌شود. یادآوری می‌شود که ما برای پر کردن مقادیر احتمال از مد ویژگی‌ها استفاده کردیم. بنابراین در توجیه علت بهبود الگوریتم با تاثیر دادن متغیر پنجم می‌توان گفت که چون اکثر خانه‌های این ستون خالی است بنابراین اکثر خانه‌های این ستون با مد داده‌ها پر می‌شود. از طرف دیگر با توجه به این که برچسب اکثر داده‌ها ۲ است بنابراین الگوریتم می‌تواند با ایجاد یک همبستگی بین این ستون با مقدار برچسب نهایی پیش‌بینی مناسبی را ارائه کند. البته این همبستگی هنگامی که مدل تنها با استفاده از مقدار ویژگی پنجم پیش‌بینی می‌کند چندان برقرار نمی‌شود چرا که در گام‌های ابتدایی الگوریتم (که گام‌های کلیدی هستند) حالت شماره یک اعداد بزرگ‌تری را به ازای خروجی‌های ۱ و ۲ متغیر پنجم نسبت می‌دهد که در نتیجه الگوریتم همواره حالت یک را پیش‌بینی می‌کند. این حالت با اضافه کردن ویژگی‌های دیگر از بین رفته و در نتیجه عملکرد الگوریتم بهتر می‌شود.

## قسمت ژ

مدل دیگری که مورد بررسی قرار گرفته است استفاده از توزیع احتمال گاوسی برای محاسبه احتمال است. در این حالت نیز احتمال ویژگی‌ها مستقل از هم فرض

جدول بررسی تاثیر وجود و عدم وجود هر یک از ویژگی‌ها: Table 2:

features involved	accuracy
(0,)	0.758
(1,)	0.389
(2,)	0.595
(3,)	0.666
(4,)	0.481
(0, 1)	0.776
(0, 2)	0.766
(0, 3)	0.762
(0, 4)	0.761
(1, 2)	0.618
(1, 3)	0.632
(1, 4)	0.517
(2, 3)	0.624
(2, 4)	0.695
(3, 4)	0.671
(0, 1, 2)	0.787
(0, 1, 3)	0.785
(0, 1, 4)	0.792
(0, 2, 3)	0.772
(0, 2, 4)	0.778
(0, 3, 4)	0.780
(1, 2, 3)	0.658
(1, 2, 4)	0.698
(1, 3, 4)	0.687
(2, 3, 4)	0.682
(0, 1, 2, 3)	0.776
(0, 1, 2, 4)	0.802
(0, 1, 3, 4)	0.794
(0, 2, 3, 4)	0.782
(1, 2, 3, 4)	0.704
(0, 1, 2, 3, 4)	0.797



شده‌اند. عملکرد مدل در این حالت در جدول ۳ آورده شده است. با مقایسه نتایج این جدول با جدول ۲ مشخص می‌شود که در حالتی که با استفاده از شمارش حالت‌ها پیش برویم عملکرد الگوریتم نسبت به حالت فعلی بهتر است. دلیل این اتفاق آن است که این مدل نسبت به مدل‌هایی که با استفاده از شمارش ساخته شده است پیچیده‌تر بوده بدین معنی که احتمال زیادی را به حالت‌هایی که تا به حال مشاهده نشده است نسبت می‌دهد در نتیجه این کار نمی‌تواند به دقت‌های خوبی روی مجموعه داده آزمون برسد.

این آزمایش با استفاده از مدل در هنگامی که ویژگی‌ها مستقل از هم فرض نشده‌اند نیز انجام شده است. در این حالت نیز عملکرد مدل نسبت به حالت عادی افت کرده است. این اتفاق دلیل مشابهی با آن چه که در پاراگراف قبلی توضیح داده شد دارد. در این حالت نیز احتمال‌های خروجی‌ها در فضای حالات پخش‌تر هستند که در نتیجه به مقادیری که واقعا اتفاق افتاده‌اند اعدادی نسبت داده نمی‌شود که این احتمال‌ها را از هم تفکیک کند. به همین دلیل مدل در طول پیش‌بینی خود معمولا تنها یک حالت را پیش‌بینی می‌کند؛ اگر حالت اولیه پیش‌بینی شده یک باشد با ادامه این پیش‌بینی می‌تواند به دقت‌های بالا برسد در غیر این صورت مدل نمی‌تواند به دقت‌های مناسبی برسد. جدول عملکرد مدل در این حالت در جدول ۴ آورده شده است.

Table 3: بررسی عملکرد مدل در حالت استفاده از احتمال گاوسی

feature	accuracy
(0,)	0.7191
(1,)	0.4016
(2,)	0.5987
(3,)	0.6616
(4,)	0.4817
(0, 1)	0.7357
(0, 2)	0.7332
(0, 3)	0.7404
(0, 4)	0.6446
(1, 2)	0.6656
(1, 3)	0.6292
(1, 4)	0.4801
(2, 3)	0.6262
(2, 4)	0.6223
(3, 4)	0.6107
(0, 1, 2)	0.7241
(0, 1, 3)	0.7567
(0, 1, 4)	0.6223
(0, 2, 3)	0.7325
(0, 2, 4)	0.6803
(0, 3, 4)	0.6958
(1, 2, 3)	0.6310
(1, 2, 4)	0.6090
(1, 3, 4)	0.5950
(2, 3, 4)	0.6653
(0, 1, 2, 3)	0.7220
(0, 1, 2, 4)	0.6734
(0, 1, 3, 4)	0.6748
(0, 2, 3, 4)	0.6915
(1, 2, 3, 4)	0.6651
(0, 1, 2, 3, 4)	0.6845

Table 4: بررسی عملکرد مدل در حالت استفاده از احتمال گاوسی بدون فرض استقلال متغیرها

features	accuracy
(0,)	0.7191
(1,)	0.4016
(2,)	0.5987
(3,)	0.6616
(4,)	0.4817
(0, 1)	0.7605
(0, 2)	0.7621
(0, 3)	0.7541
(0, 4)	0.7722
(1, 2)	0.4948
(1, 3)	0.4786
(1, 4)	0.4109
(2, 3)	0.5203
(2, 4)	0.3784
(3, 4)	0.4033
(0, 1, 2)	0.7622
(0, 1, 3)	0.7458
(0, 1, 4)	0.7456
(0, 2, 3)	0.7846
(0, 2, 4)	0.7522
(0, 3, 4)	0.7339
(1, 2, 3)	0.3672
(1, 2, 4)	0.3527
(1, 3, 4)	0.3545
(2, 3, 4)	0.3417
(0, 1, 2, 3)	0.7670
(0, 1, 2, 4)	0.7606
(0, 1, 3, 4)	0.7419
(0, 2, 3, 4)	0.7290
(1, 2, 3, 4)	0.3333
(0, 1, 2, 3, 4)	0.7429