به نام خدا



دانشگاه صنعتی امیرکبیر

(پلی‌تکنیک تهران)

درس شناسایی آماری الگو

استاد رحمتی

تمرین اول

علیرضا مازوچی

۴۰۰۱۳۱۰۷۵

سوال ۱

پیش از جواب‌دادن به سوالات لازم است تا روش طراحی سامانه را تبیین کنم. من قصد دارم تا سه دسته‌بند طراحی کنم و سپس این سه دسته‌بند در یک چهارچوب گروهی (Ensemble) جواب نهایی را برگردانند. دسته‌بند اول با چهره شخص و یک شبکه عصبی تشخیص می‌دهد که فرد با چه احتمالی در بازه مناسب قرار دارد. دسته‌بند دوم با پردازش صوت کودک با یک شبکه عصبی دیگر تشخیص می‌دهد که کودک به چه احتمالی در بازه سنی مناسب قرار دارد یا نه؟ دسته‌بند سوم با ویژگی قد و وزن کودک احتمال تعلق داشتن کودک به بازه سنی مناسب را بیان می‌کند. نهایتا خروجی هر سه دسته‌بند با به یک MLP‌ ساده می‌دهیم تا نتیجه نهایی را اعلام کند.

الف) مسئله از نوع دسته‌بندی یا Classification است؛ چراکه قرار است کل بازدیدکنندگان را به دو کلاس گروه سنی مناسب و گروه سنی نامناسب تقسیم کنیم.

ب) به سنسور تشخیص چهره،‌ صوت، قد و وزن نیاز داریم.

ج) ما یک مجموعه‌داده نیاز داریم که برای هر داده یک تصویر از تمام‌رخ چهره، یک صوت بعد از خواندن یک متن، قد و وزن را در خود داشته باشد.

د) می‌توان داده‌ها را از بازدیدکنندگان فعلی تا زمانی استقرار سامانه تهیه کرد. یعنی هر تصویر و سایر مشخصات هر کودک بازدیدکننده را می‌توان برداشت و یک ناظر انسانی مسئول تعیین کلاس کودک می‌شود. باتوجه به آمار بالای بازدید پارک دیزنی، با این شیوه می‌توان در مدت زمان قابل قبولی حجم مناسب داده جمع‌آوری شود. البته باید توجه داشت ممکن است قوانین حفظ حریم خصوصی مانع از جمع‌آوری داده تمام بازدیدکننده‌ها بدون اجازه آن‌ها باشد که در این صورت لازم است این جمع‌آوری داوطلبانه و با آگاهی بازدیدکننده‌ها صورت بگیرد.

علاوه بر شیوه فوق می‌توان از مجموعه‌داده‌های آماده هم کمک گرفت که البته هر کدام بخشی از ویژگی‌ها مدنظر ما را دارند و می‌توانند در توسعه زیرمدل‌های ما مورد استفاده قرار بگیرند.

ه) ویژگی قد، وزن،‌ یک تصویر که یک آرایه دوبعدی از ویژگی‌هاست و صوت که یک آرایه یک‌بعدی از ویژگی‌هاست.

و) برای ویژگی قد و وزن پیش‌پردازش خاصی نیاز نیست. در مورد صوت، لازم است تا پیش‌پردازشی برای حذف نویز محیط و سایر نویز‌های این حوزه انجام گیرد. در مورد تصویر هم باید پیش‌پردازش‌هایی برای شفاف کردن عکس، حذف لرزش موقع تصویر و تاری عکس و مواردی از این دست صورت بگیرد.

ز) مهم‌ترین چالش جمع‌آوری داده است. باتوجه به پیچیدگی مدل و دارابودن انواع ویژگی‌ها نیاز به تعداد داده بالاست. قوانین حفظ حریم خصوصی هم ممکن است کار را سخت‌تر بکنند. در این حالت ممکن است داده به اندازه جمع‌آوری نشود که مجبور خواهیم بود از مدل‌های ساده برای بخش تصویر و صوت استفاده کنیم که بی‌شک در دقت مدل موثر است. حتی این احتمال وجود دارد که مجبور باشیم بخشی از مدل را مثلا بخش تصویر را حذف کنیم.

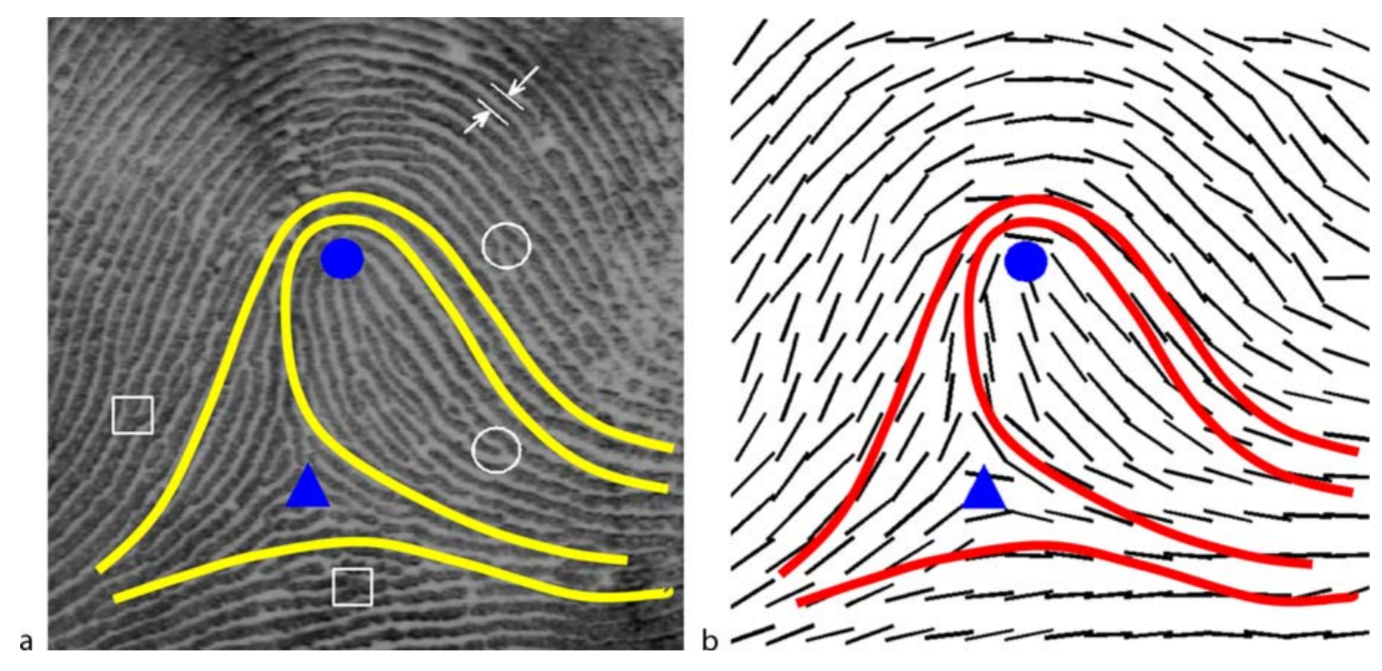
یکی دیگر از چالش‌های سیستم جدید،‌ امکان تغلب در آن است. درحالی که ویژگی‌هایی مانند قد به سختی مورد تغلب واقع می‌شود ولی در مورد صوت و تصویر ممکن است کاربر تغییراتی در صدا و چهره خود بدهد که سیستم را به شک بیاندازد. در مورد وزن هم امکان افزایش آن با لباس‌های سنگین وجود دارد. همچنین به دلیل کرونا ( و بیماری‌های آتی!) ممکن است امکان دریافت چهره کامل به سختی صورت بگیرد.

ح) نقطه قوت اصلی سیستم، دقت بیشتر و توجه به جوانب مختلف نسبت به سیستم قبلی است. در سیستم قبلی تنها به قد توجه می‌شد که ممکن بود یک کودک که قد آن نسبت سن واقعی‌اش متداول نباشد را به خطا بیاندازد درحالی که در این سیستم مبتنی بر چندین زیرمدل است. نقطه قوت دیگری که شاید بتوان برای آن درنظر گرفت آن است که خروجی سیستم فعلی برای کودکان و والدینشان پذیرفتنی‌تر است. در سیستم قبلی محتمل بود تا برخی از کودکان و والدین توجه تنها به یک خط‌کش را عادلانه ندانند درحالی که در سیستم فعلی و پس از جمع‌آوری اطلاعات مختلف و نهایتا تصمیم‌گیری توسط کامپیوتر امکان متقاعدکردن بهتری وجود خواهد داشت.

درکنار نقاط قوت، نقاط ضعفی وجود دارد. سیستم فعلی نسبت به سیستم قبل، پیچده‌تر است و برای استفاده از آن هم نیاز به تجهیزات مختلف سخت‌افزاری و نرم‌افزاری است. استفاده از سیستم پیشنهادی وقت کاربران را می‌گیرد چراکه آن‌ها باید چهار تست را به جای یک تست بدهند. اگرچه می‌توان این‌ها را باهم گرفت. (مثلا در حین ایستادن بر روی وزنه، قد و تصویر هم برداشته شود و در همین حین از کودک خواسته شود یک متن یا شعر را بخواند.) همچنین جمع‌آوری داده‌ها همانطور که گفته شد سخت، زمان‌بر و پرهزینه است.

سوال ۲

الف) می‌توان محل تورفتگی‌ها را مانند یک خط به یکدیگر وصل کرد. سپس از الگو‌های این خط‌ها ویژگی‌هایی را پیشنهاد داد. مثلا می‌توان محلی که این خط‌ها سه‌شاخه شده است را ذخیره کرد؛ مانند مربع‌های توخالی در شکل زیر. همچنین می‌توان محل‌هایی که خط‌ها خاتمه پیدا کردند را هم به عنوان ویژگی ذخیره کرد؛ مانند دایره‌های توخالی در شکل زیر.[[1]](#footnote-1)



ب) وضعیت لب‌ها و وضعیت ابرو می‌تواند دو ویژگی قابل محاسبه و مفید برای این مسئله باشد. توجه کنید که در این حالت لب‌ها و ابرو‌ها را مانند قسمت قبل با یک خط مدل می‌کنیم. این ویژگی‌های ساده‌ای چراکه به آسانی می‌توان دید در احساسات مختلف وضعیت لب و ابرو متفاوت است.

ج) می‌توان تصاویر متوالی از یک فرد را دریافت کرد و میزان حرکت سر و دست او را به عنوان ویژگی حساب کرد. (با اختلاف تصاویر) چراکه برخی از افراد هنگام حرکت، دست‌هایشان یا سرشان بیشتر تحرک دارد. همچنین می‌توان نسبت اندازه قسمت‌های مختلف بدن به یکدیگر در یک تصویر را هم درنظر گرفت.[[2]](#footnote-2)

د) با توجه به داشتن تضاویر متوالی از یک فرد می‌توان ویژگی توقف را برای هر شخص تعریف کرد. این ویژگی برای تشخیص فرد ایستاده متوقف، نشسته و مانند آن مناسب است. همچنین می‌توان پس از تشخیص هر فرد، نسبت ارتفاع به عرض کوچکترین مستطیلی که شخص در آن قرار گرفته است را هم بررسی کرد. این ویژگی هم کمک می‌کند متوجه شویم فرد روی پا ایستاده یا خیر.

ه) تعداد گوشه (تعداد شکستگی)، نسبت پیکسل‌های الگو در حدود بالاترین (یا پایین‌ترین،‌ یا راست‌ترین یا چپ‌ترین) قسمت الگو نسبت به کل پیکسل‌ها، نسبت مساحت محصور در الگو به محیط الگو

و) تعداد گوشه (تعداد شکستگی)، نسبت مجموع پیکسل‌های الگو در حدود بالاترین و پایین‌ترین (یا مجموع راست‌ترین و چپ‌ترین) قسمت الگو نسبت به کل پیکسل‌ها، نسبت مساحت محصور در الگو به محیط الگو

ز) تعداد گوشه (تعداد شکستگی)

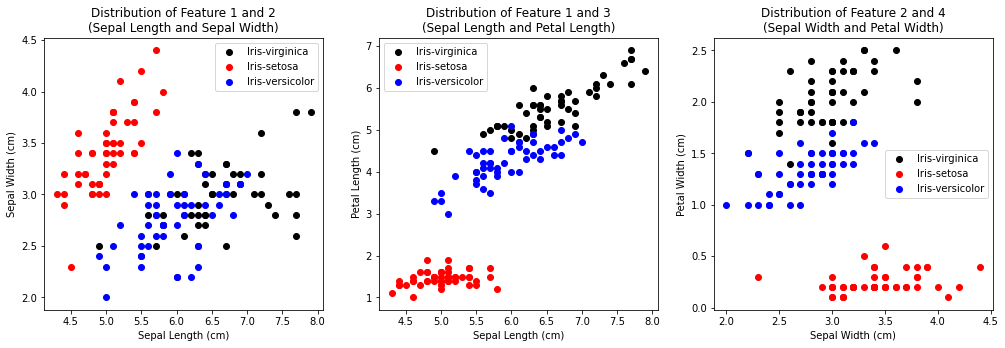
ح) تعداد گوشه (تعداد شکستگی)، نسبت پیکسل‌های الگو در حدود بالاترین (یا پایین‌ترین،‌ یا راست‌ترین یا چپ‌ترین) قسمت الگو نسبت به کل پیکسل‌های الگو، نسبت مساحت محصور در الگو به محیط الگو

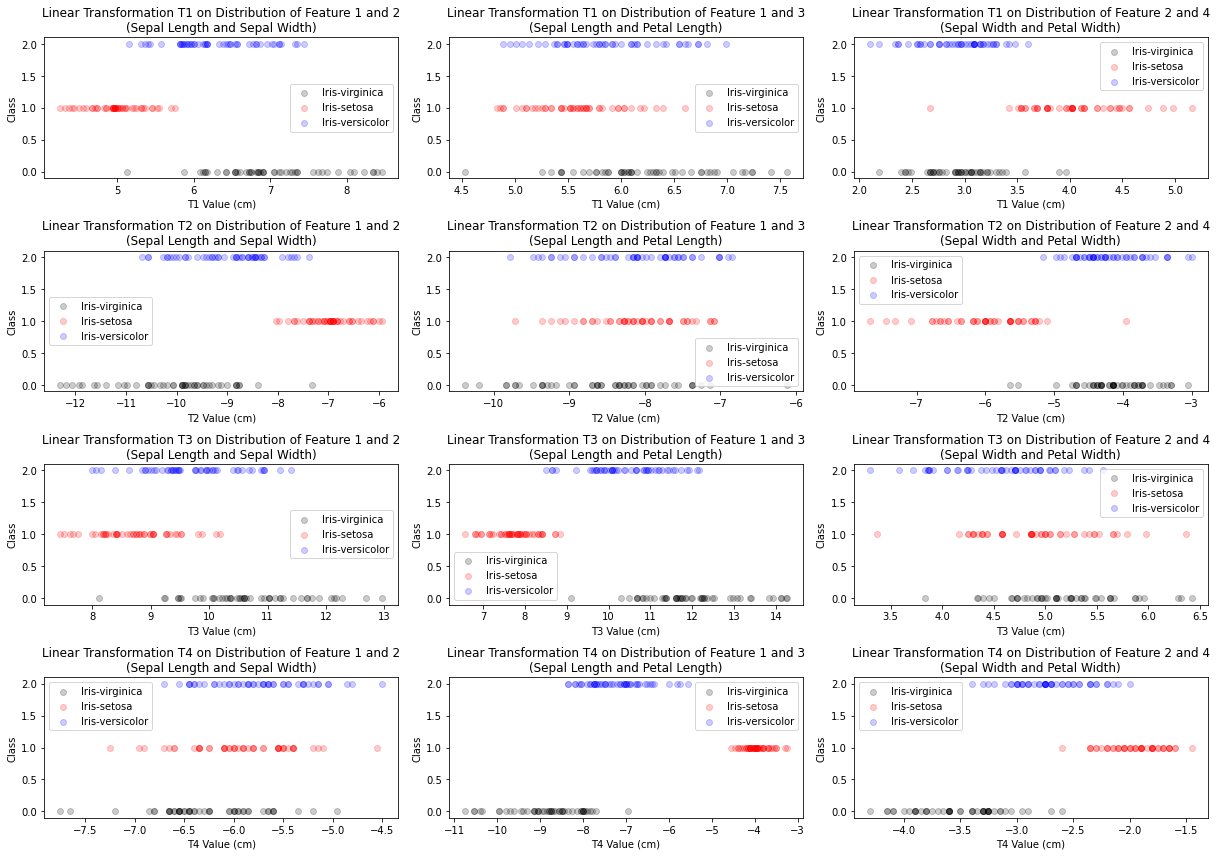
ط) تعداد گوشه (تعداد شکستگی)، نسبت مساحت محصور در الگو به محیط الگو

ی) تعداد گوشه (تعداد شکستگی)، نسبت مساحت محصور در الگو به محیط الگو، مجموع پیکسل‌ها در بالاترین، پایین‌ترین، راست‌ترین و چپ‌ترین قسمت الگو نسبت به کل پیکسل‌های الگو (‌به شرط عدم چرخش مربع و مثلث)

سوال ۳

الف)



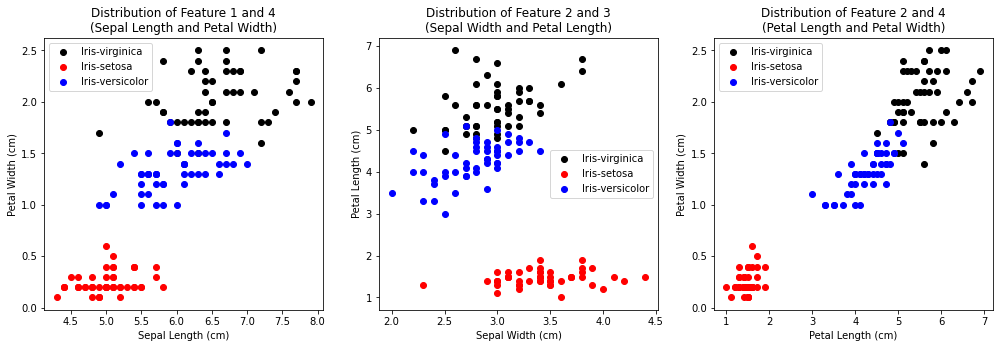
ب)

ج) برای ویژگی ۱ و ۲، از همه مناسب‌تر به نظر می‌رسد اگرچه هم قابل قبول است. دو تبدیل و نتوانسته‌اند برخلاف دو تبدیل دیگر، کلاس نسبتا مجزای قرمز را به خوبی جداکنند. در مورد و هر دو تقریبا مشابه هستند اما تبدیل کمی کلاس قرمز را بهتر جدا کرده است.

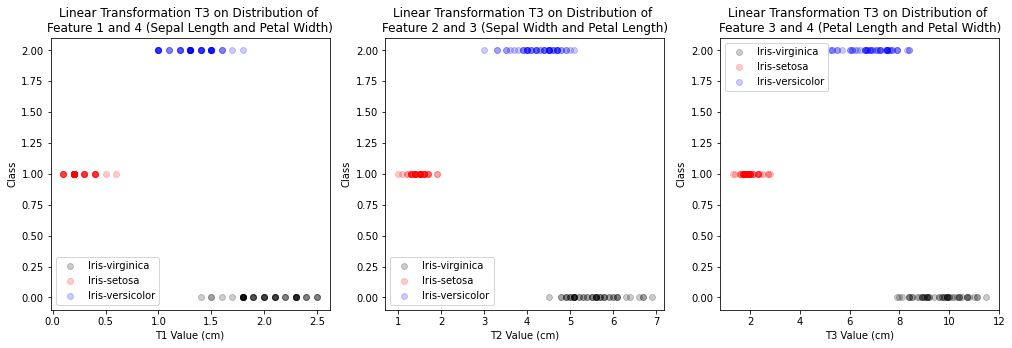
برای ویژگی ۱ و ۳، مناسب‌ترین است. در تبدیل و داده‌های هر سه کلاس روی هم افتاده است. در تبدیل هم دو کلاس آبی و سیاه همپوشانی زیادی دارند و داده‌های کلاس قرمز نسبت به پراکنده‌تر است.

برای ویژگی ۲ و ۴، مناسب‌ترین است.‌ تبدیل داده‌های هر سه کلاس را در یک نقطه نگاشت می‌کند. و دو کلاس آبی و سیاه را روی هم می‌اندازد درحالی که در سه کلاس تاحد قابل قبولی از هم جدا شده‌اند.

د)



ه) با توجه به نمودار‌های قسمت د می‌توان حدس زد که برای جفت ویژگی (۱ و ۴) و (۲ و ۳) خطوط موازی با محور x‌ها با دقت بسیار خوبی می‌توانند داده‌های کلاس‌های مختلف را از هم جدا کنند. لذا برای این دو قسمت از تبدیل خطی ]۱ و ۰ [ استفاده کردم. برای جفت ویژگی ( ۲ و ۴ ) شاید خطوط مورب 45 درجه ( بدون توجه به مقیاس) به یکی از بهترین نتایج ممکن ختم شود. نهایتا با اختلاف مقیاس دوبرابر ویژگی اول و دوم، از تبدیل خطی ]۲ و ۱ [ استفاده کردم:



سوال ۴

الف) پرتاب سکه از توزیع دوجمله‌ای پیروی می‌کند.

ب) برای قسمت اول این سوال باید سه توپ از شش توپ را انتخاب کرد به عنوان شوت موفق و سپس باید سه توپ با احتمال موفقیت گل شوند و سه توپ دیگر با احتمال شکست به بیرون بروند. برای قستم دوم هم باید شوت اول و دوم بیرون برود و شوت سوم حتما باید گل شود. مابقی شوت‌ها اهمیتی ندارد.

ج)

این معادله را من به صورت دستی نتوانستم حل کنم. با سپردن این معادله به ابزار‌های آنلاین متوجه شدم این معادله تنها یک جواب حقیقی دارد و آن در حدود 2.847 بوده است.

برای قسمت بعدی سوال از جدول مقادیر تجمعی این توزیع[[3]](#footnote-3) و استفاده کردم.

د)

برای قسمت چهارم این سوال می‌توانیم با یک مثال نقض نشان دهیم که ممکن است دو ویژگی X و Y همبستگی نداشته باشند ولی با یکدیگر رابطه داشته باشند.

در این مثال X و Y‌ با یکدیگر رابطه دارند ولی همبستگی‌شان برابر با صفر است!

سوال ۵

الف)

طبق فرض سوال برابر با است که اثبات می‌کند یک بردار ویژه است و مقدار ویژه آن برابر با است.

ب)

به طور مشابه در می‌یابیم نیز یک بردار ویژه است و مقدار آن برابر با یا است.

ج)

با این رابطه می‌توان دو مقدار ویژه را پیدا کرد.

در رابطه فوق برابر بودن با با استفاده از و برقرار خواهد بود و نهایتا متناظر با این دو مقدار ویژه،‌ بردار‌های ویژه زیر را خواهیم داشت:

این رابطه در شرایطی برقرار است که . حال اگر باشد ولی باشد می‌توان نوشت:

اما اگر باشد باید به نحو دیگری عمل کرد:

اگر باشد یک عبارت صفر می‌شود لذا بردار‌های ویژه برابر خواهد بود با:

د)

ه)

به نظر می‌رسد هیچ وضعیتی غیرصفری وجود ندارد که ثابت بماند.

و) اگر ‌ باشد لذا برای مقدار خواهیم داشت ( از روی مقدار و بردار ویژه هم می‌توانستیم این را بفهمیم):

با توجه به اینکه برابر ضریبی از شد می‌توان نتیجه گرفت:

لذا داستان خسرو و شیرین به این شکل خواهد شد که در ابتدا خسرو مقداری شیرین را دوست دارد و شیرین به همان اندازه از او متنفر است (یا بالعکس). سپس با گذشت هر لحظه این دوستی و تنفر به نصف مقدار قبل کاهش می‌یابد تا سرانجام هیچ حسی میان این دو نفر وجود نداشته باشد.

ز)

به نظر می‌رسد که هر دو بعد آرایه در حال نزدیک‌شدن به ۴ هستند. می‌توان این مورد را دقیق‌تر بررسی کرد:

پس در این حالت، در ابتدا خسرو 3 واحد به شیرین علاقه‌مند است و شیرین 5 واحد به خسرو و به گذشت زمان محبت این دو شخص باهم یکسان می‌‌شود و هر دو ۴ واحد به یکدیگر علاقه‌مند خواهند بود. همچنین اختلاف علاقه‌شان به یکدیگر در هر گام نصف خواهد شد.

ح)

ط) این حالت هم مشابه قسمت‌های قبل انتظار می‌رود با ضرب‌های A در بردار ویژه به همان بردار ویژه با مقدار ویژه متناظر با آن برسیم:

ی)

لذا در این حالت دوستی خسرو و شیرین هربار دو برابر می‌شود و به بی‌نهایت میل پیدا می‌کند!

ک)

ل)

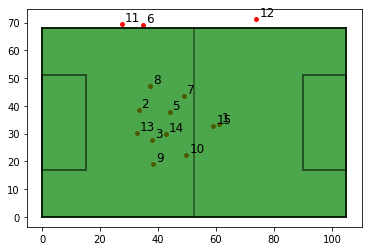
در این مورد اگر در ابتدا خسرو شیرین را دوست داشته باشد و شیرین به همان میزان از خسرو بدش بیاید، در هر بار هم علاقه و هم نفرت سه برابر می‌شود تا در نهایت خسرو بی‌نهایت شیرین را دوست داشته باشد و شیرین بی‌نهایت از خسرو متنفر شود! به طور مشابه اگر در ابتدا شیرین خسرو را مقداری دوست داشته باشد و به همان میزان خسرو از شیرین متنفر باشد در نهایت شیرین بی‌نهایت به خسرو علاقه‌مند خواهد بود درحالی که خسرو بی‌نهایت از شیرین متنفر خواهد بود. اگر در ابتدا هر دو هیچ‌حسی نسبت به همدیگر نداشته باشند تا ابد و در هرلحظه این بی‌حسی ادامه‌دار خواهد بود.

م)

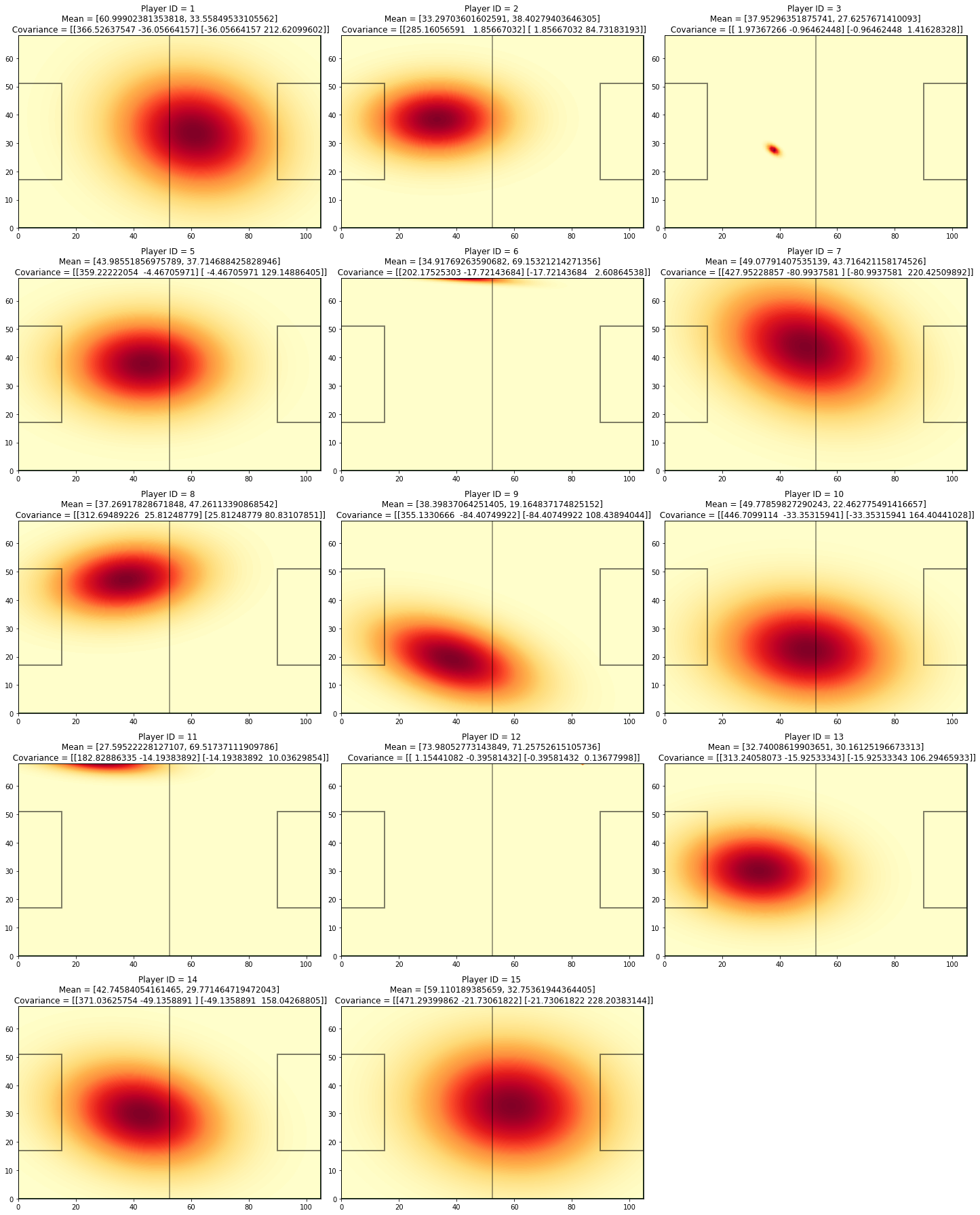
در این حالت در ابتدا خسرو و شیرین همدیگر را به یک میزان دوست دارند. در لحظه بعد به همان میزان از یکدگیر بدشان می‌آید و مجددا باهم دوست می‌شوند و تا انتها یک لحظه در میان باهم دوست می‌شوند و قهر می‌کنند!

سوال ۶

الف) در فایل موجود چهارده بازیکن دیده می‌شوند. به نظر سه بازیکن 6، ۱۱ و 12 بازیکنان ذخیره هستند ولی با این حال ترسیم شده اند.



ب)

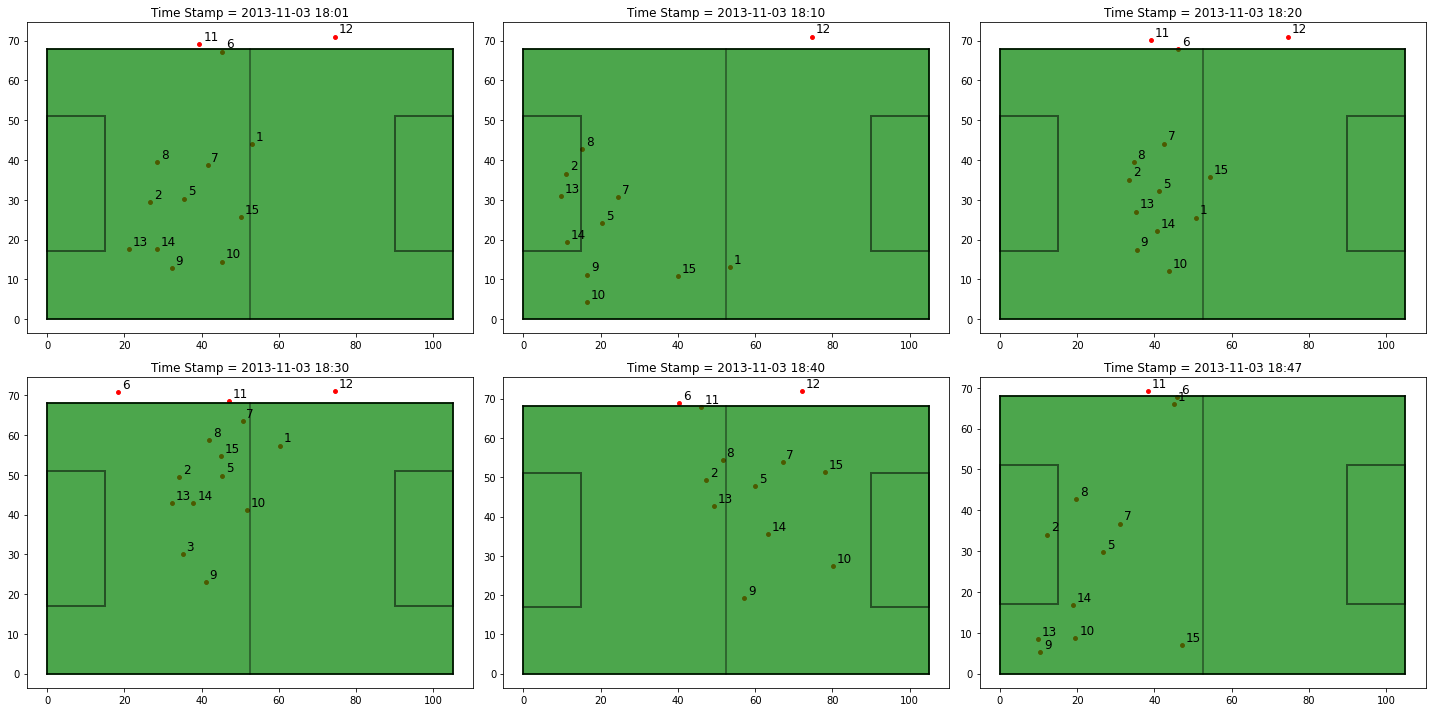


ج)



د) برای تحلیل نقش بازیکنان، علاوه بر مراجعه به نمودار‌های قسمت قبل، چندین تصویر در لحظات مختلف بازی را هم ترسیم کردیم. به علاوه به انرژی و فاصله پیموده‌شده نهایی توجه داشتیم.

بازیکن شماره 6، 11 و 12 باتوجه به میانگین موقعیت و مجموع فاصله طی کرده‌شان بازیکنان ذخیره بودند. بازیکن شماره ۳ ظاهرا بازی نکرده است یا شاید هم برای آن داده زیادی جمع نشده است. هم اینکه تنها از دقیقه 26 تا 32 برایش داده وجود دارد و هم اینکه در کل این مدت در یک محل قرار داشته است. با کنار گذاشتن این بازیکنان، ده بازیکن باقی می‌مانند که باتوجه به نمودار‌ها، احتمالا دروازه‌بان ترسیم نشده است؛ قسمت چپ زمین طرف خودی است و سیستم بازی به شکل 2-4-4 است. در جدول زیر می‌توانید نقش‌های پیشنهادی من را مشاهده بفرمایید.



|  |  |
| --- | --- |
| شماره بازیکن | نقش پیشنهادی |
| ۱ | مهاجم |
| ۲ | دفاع وسط |
| ۳ | ؟! |
| ۵ | هافبک وسط |
| ۶ | ذخیره |
| ۷ | هافبک چپ |
| ۸ | دفاع چپ |
| ۹ | دفاع راست |
| 10 | هافبک راست |
| 11 | ذخیره |
| 12 | ذخیره |
| 13 | دفاع وسط |
| 14 | هافبک وسط |
| 15 | مهاجم |

سوال ۷[[4]](#footnote-4)[[5]](#footnote-5)[[6]](#footnote-6)

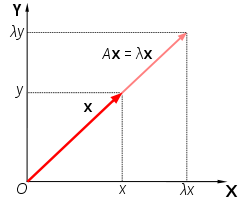
الف) Whitening Transform یک تبدیل خطی است که آرایه‌ای از متغیر‌های تصادفی با یک ماتریس کواریانس را به آرایه‌ای از متغیر‌های تصادفی تبدیل می‌کند که دیگر بین آن‌ها هیچ همبستگی وجود ندارد و واریانس هر متغیر تصادفی هم برابر با یک خواهد شد. پس با درنظر گرفتن این تبدیل به عنوان یک گام پیش‌پردازش می‌توان متغیر‌های تصادفی را از یکدگیر جدا کرد. چنین چیزی باعث می‌شود در برخی از الگوریتم‌های یادگیری ماشین آموزش بهتری داشته باشیم. (به همان دلیل که ویژگی‌های هم‌بسته زائد است!)

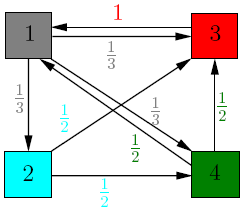
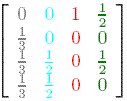
ب) شاید بتوان گفت فضای پیوسته حالت عام فضای گسسته است! رگرسیون را می‌توان حالت کلی‌تری از دسته‌بندی دانست چون می‌توان برای هر کلاس یک نقطه در فضای پیوسته درنظر گرفت؛ لذا در این صورت یک الگوریتم رگرسیون می‌تواند هر داده را به نزدیک‌ترین یکی از نقطه‌های متناظر با کلاس‌ها در فضای پیوسته و کلاس آن نسبت دهد. یا شاید بتوان برای هر کلاس یک بعد پیوسته در نظر گرفت و احتمال تعلق به آن کلاس را با آن بعد نشان داد. طبیعتا هر داده به کلاسی تعلق دارد که مقدار آن بعدش بیشتر باشد. و سایر ایده‌های مشابه...

ج) یک بردار ویژه برای یک ماتریس M، تبدیلی غیرصفر است که اگر بر ماتریسM اعمال شود حاصلش نسبتی از همان ماتریس M شود. به آن نسبت هم، مقدار ویژه متناظر با آن بردار ویژه گفته می‌شود.

با تعریف رسمی‌تر می‌توان گفت که اگر M یک تبدیل خطی از فضای برداری V به خودش در میدان F باشد و v یک بردار غیرصفر در فضای v‌ باشد؛ آنگاه v یک بردار ویژه برای M است اگر یک ضریب اسکالر از v باشد که می‌توان آن را با فرمول زیر نوشت. در این فرمول یک عدد اسکالر در میدان F است که مقدار ویژه متناسب با بردار ویژه v درنظر گرفته خواهد شد.

د) از نظر هندسی، یک بردار ویژه را می‌توان جهتی دانست که اگر یک تبدیل خطی بر آن اعمال شود و آن را بکشد، همچنان در همان جهت یا خلاف آن جهت باقی بماند و چرخش نداشته باشد. مقدار ویژه مقیاسی است که مشخص می‌کند با این تبدیل به چه میزان بردار ویژه کشیده شده است. همچنین منفی‌بودن مقدار ویژه نشان‌دهنده آن است که پس از اعمال تبدیل خطی، بردار ویژه در خلاف جهت اولیه قرار می‌گیرد. به عنوان مثال در تصویر زیر مشخص است که اگر x بردار ویژه تبدیل A‌ باشد. پس از اعمال A‌ بر آن، همچنان جهت x‌ تغییر نکرده است. فقط مقدار آن عوش شده است که نسبت آن وابسته به مقدار ویژه یعنی است.



ه) گوگل برای آنکه صفحات را رتبه‌بندی کند از الگوریتم PageRank استفاده می‌کند. به طور خلاصه، متناسب با این الگوریتم اعتبار هر صفحه متناسب با اعتبار صفحاتی است که به آن ارجاع داشته‌اند. در این الگوریتم یک ماتریس N\*N‌ به نام A‌ ساخته می‌شود که N تعداد صفحات است. در ماتریس A مشخص می‌شود که یک صفحه به یک صفحه دیگر ارجاع داده است یا خیر. نمونه‌ای از این ماتریس برای یک گراف آورده شده است:

حال باید امتیازها را به گونه‌ای به ماتریس به گره‌ها تخصیص داد که روابط برقرار بماند یعنی امتیاز هر راس برابر باشد با امتیاز گره‌هایی که به آن ارجاع داده‌اند ضرب در یک عدد نرمال‌سازی. لذا اگر برداری از امتیاز رئوس را v بمانیم باید داشته باشیم:

برای پیدا کردن v‌ یک راه آن است که ابتدا مقدار تصادفی برای آن درنظر بگیریم و سپس به تعداد بار زیاد A‌ را بر آن اعمال کنیم تا به یک بردار همگرا شود و یک راه دیگر آن است که بردار ویژه متناسب با مقدار ویژه ۱ برای ماتریس A‌ را پیدا کنیم که روال سریع‌تری محسوب می‌شود.

1. https://sci-hub.se/10.1007/978-0-387-73003-5\_50 [↑](#footnote-ref-1)
2. https://sci-hub.se/10.1117/12.2015634 [↑](#footnote-ref-2)
3. https://www.york.ac.uk/depts/maths/tables/poisson.pdf [↑](#footnote-ref-3)
4. https://en.wikipedia.org/wiki/Eigenvalues\_and\_eigenvectors [↑](#footnote-ref-4)
5. https://www.dhruvonmath.com/2019/03/20/pagerank/ [↑](#footnote-ref-5)
6. http://pi.math.cornell.edu/~mec/Winter2009/RalucaRemus/Lecture3/lecture3.html [↑](#footnote-ref-6)