

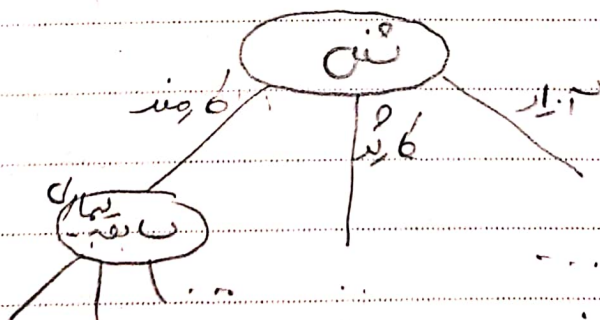
دعوت ۱: تسک های که توسط علم داده طی قابل انجام است عبارتند از: خوشه بندی ها ۱
 دسته بندی داده ها یا ایجاد مدل های برای پیش بینی، کشف association rule ها ۲
 میان آشی ها یا attri ها یا داده ها یا تشخیص نوین در داده ها یا یک سری داده ها حذف ۳
 داده حذف قطع و مضاعف است. ۴

۱۱) کشف association Rules: هر توانی اشک های تکلیلی که در داده رخ می دهد را تشخیص ۵
 دارد. مثلاً هر توانی تشخیص داد افرادی که مثلاً باقیه بیماری منشی رخن داشتند اکنون ۶
 برای بیماری قلبی مشابه نشانه داده اند که هر توانی با محاسبه Confidence و support ۷
 برای attri این اشک ها را ثبت می کند. مثلاً اگر $Confidence = 0.7$ و $Support = 0.3$ ۸
 یعنی که مثلاً بیماری قلبی رخن با احتمال ۰.۷ به آ برای بیماری قلبی منشی رخن خواهد بود ۹
 و اگر $Support = 0.3$ و یعنی تمام داده های رخن در دیتا هستند که بیماری ۱۰
 قلبی رخن و بیماری قلبی با هم وجود داشته اند. ۱۱

مثلاً هر توانی کشف کرد افرادی که شغل خاص دارند به احتمال ۰.۷ به فلان بیماری مبتلا ۱۲
 مثلاً هر توانی. ۱۳

مثلاً افرادی که مبتلای به فلان بیماری است به احتمال ۰.۷ به یک بیماری خاص مبتلا ۱۴

۱۲) ایجاد مدل های پیشگانه: گاهی بندی فرایند یافتن یک مدل است که گاهی داده ها ۱۵
 از هم جدا می کنند. مدل بیکیه تحلیل داده ها را آموزش می دهد و عمل می کند. از مدل برای پیش بینی ۱۶
 به سبب گاهی نمونه داده های بیشتر استفاده می شود. یک روش برای ایجاد مدل از نمونه ۱۷
 از دیتا تقسیم است. هر توانی با این روش تشخیص دهیم فرایند کشف داده ها ۱۸
 از بیکیه از داده ها به سبب تا سبب به سبب به سبب به سبب به سبب به سبب به سبب به سبب ۱۹
 هر توانی مثلاً attri شغل را در فواصل درخت قرار داد. ۲۰



Arman

بیماری دفعه دوم

بیماری کم خونی

(۳۱) خودی شدن داده ها، خودی بینی آبجکت ها، اتصالات و در خودی ها، مختلف قرار
 می دهد. به عنوان اینکه توصیف به بدیستب ها، اولی که این ها داشته باشد.
 ۳ در مواردی که بدیستب ها، کلاس ها، خاصیت نسبت یا وضعیت بیمار معوض نسبت
 می تواند باز هم بیمار از این طریق اصولی در خودی ها، مختلف و متغیر قرار دارد و بیان
 نسبتی می تواند در این ... نسبت عمل کند.
 ۴ این در واقع باید یک نسبت خاص باشد و نسبت به این نسبت که در این نسبت
 نسبتی است قرار گیرند. پس این روش بیمار از این که وضعیت نسبت به این نسبت خودی
 ۸ قرار می دهد و به همین نسبت دیگر بیمار از این نسبت عمل کند.

(۳۲) Anomaly detection: ممکن است در بیمار از منابع گفته شده خودی ها و وجود
 ۱۱ نسبتی باشد که مطابق با رفتار مایل نگه داده نباشند. این افراد داده ها، نسبت یا نسبت
 ۱۲ نسبت به حساب می آیند که می توان این میزان را نسبت به داده ها و نسبت افراد جدا
 ۱۳ کنیم و نسبت به این نسبت ها، نسبت می کنیم. مثلاً ممکن است فردی با یک بیمار، تادر منابع
 ۱۴ کند و مطابق با این نسبت نگه داده ها، نسبت می شود که با این فرد، نسبت خاص در نظر گرفته
 ۱۵ و به همین نسبت دیگر بیمار از این نسبت می تواند استفاده از نسبت ها، داده ها و این
 ۱۶ نسبت را بدست آورد.

فصل ۲

۲۰ **نسبت بدیستب**، نسبتی است که این نسبت به مقدار بیان آن می تواند وجود
 ۲۱ داشته باشد و گفته وسیع تر دارد.

۲۳ نسبت است. چون بدیستب نسبت بیان می دهد و می تواند گفته شود ۱۱ بیان دیگر
 ۲۴ که مقایسه نسبت آن معنای ندارد و می تواند فقط به مقدار عددی در این بیان
 ۲۵ داشته باشد.

۲۷ ratio است. چون مقدار صفت آن واقعی و معنی را است. مثلاً نوزاد تازه متولد شده
 Arman. این مقدار در این معیار متغیر می شود.

روشنایی یا نور منبع انوار رنگی در λ بودن یا نباشد نیست. چون بین اوجات یا مقدار
بازده می تواند بگیرد و گستره وسیع تری دارد.

بیولم است. چون می تواند مقدار دقیق داشته باشد و فقط از عمل اعداد صحیح
منتهی شد مثلاً روشنایی می تواند $\frac{3}{4}$ باشد.

ratio است. چون رفع روشنایی در منبع می تواند مقدار منفی انذارش کند
صرفاً منفی مقدار شدت انذارش می کند.
انحراف در رفع روشنایی می شود یعنی در نقطه منفی واقع است و به معنی عدم
توجه به هیچ گونه نذر هست.

روشنایی که با نور افراد را می رود یا نباشد نیست. چون بین اوجات یا مقدار می تواند
بیش از آن در مقدار داشته باشد. مثلاً افراد می توانند مقدوره کم، متوسط، زیاد و درجه
و شتاب بیان کنند.

گستره است. چون افراد دقیق نمی توانند یک مقدار دقیق بین آن دهند و صرفاً با بقیه بین
کم و زیاد متوسط و یا با تقویم بین یک مقدار صحیح بیان می کنند.

ordinal است. چون نور افراد می تواند به عدد ضعیف، کم، متوسط، زیاد، و غیر
بیان شود و این مقادیر دارای یک تقویم حقیقی یا بقیه بین هستند.

بازده اندازه گیری λ یا تقوله: یا نباشد نیست. چون بین اوجات یا مقدار می تواند
داشته باشد و رنگ شده $32, 44, 56$ (تا 100) متغیر است و مقدار می تواند بگیرد.

بیولم است. چون می تواند مقدار دقیق نیز بگیرد و فقط مقدار صحیح نمی گیرد.
مثلاً می تواند مقدار $37, 45$ بگیرد.

Arman

1. $ratio$ است. چون هر توان گفت 10^5 برابر ۱۰ است. در واقع نسبت
2. (برابر) ... برای آن هر توان تعریف کردیم یعنی صفر درجه یعنی دایره و
3. و نقطه صفر واقع است. علاوه بر این مقدار صفت ۱۰ به ندره و قدره همان
4. اختلاف بین مقدار میسیم. اما آنرا می‌توانیم با $interval$ نشان دهیم.

5

6. **حال حال اعداد** : با این نسبت. چون حال ها می‌توانند به نفع و
تضرر باشند و واقع به حالت برای آن ها تعریف شده در حالت.

8

9. نکته است چون مقدار می‌تواند از نفع و بدتر می‌تواند باشد. با این و مقدار بین این
10. به مقدار تعریف می‌شود و در واقع مقدار صفت می‌تواند در این بازه مقدار نسبت.

11

12. $ordi$ است. چون حال ها می‌توانند در واقع یک رتبه بندی را مشخص کنند و این
13. به حال برای یک ترتیب می‌باشد.

14

15. **ارتفع از سطح دریا** : با این نسبت چون می‌تواند یک کوه و دریا از مقدار را بگیرد
16. و صرفاً در حالت برای آن تعریف شده.

17

18. سیر است. چون ارتفاع هر توان در مقدار از این صفت بگیرد و صرفاً عدد صریح
19. نمی‌گیرد.

20

21. $ratio$ است. چون نسبت هر توان برای آن تعریف کردیم. مثلاً ارتفاع ۱۰ صفر و برابر
22. ارتفاع ۱۰ است. همچنین ارتفاع صفر یعنی دایره و صفر واقع است و همان
23. سطح دریا است.

24. اما آنکه بخواهیم ارتفاع منفی این را با این بد از سطح را گذاریم کنیم با توجه به
25. آنکه شامل مقدار منفی خواهد شد $interval$ خواهد بود.

26

27

Arman

تعدادها این یک بهار تان. با نسی نیست چون بیدار در مقدار حالت
می تواند باشد و صرفاً مقدار می شود برای آن تعریف شده است و می تواند صرفاً
افراد این یا صدها یا در بسیار شدن باشند.

گفته است. چون تعداد افراد هیچ وقت نمی تواند مقدار صفر داشته باشد و صرفاً مقدار
صفر می تواند باشد.

ratio است. چون نسبت می تواند برای تعدادی این افراد تعریف شود و تعداد آن
مثلاً دو برابر تعداد آن باشد و این مقدار معنی دارد و فقط صفر است.
انقضی در این مقدار نیست به بود و در به محاسبه افتلاف مابین این مقدار می باشد.
تعداد تعداد نمی تواند صفر باشد.

ISBN: با نسی نیست. چون یک کد مشخصه فرید برای هر کتاب منتشر شده می باشد
بین این دو مقدار تفاوت است.

گفته است. یک کد ۱۳ رقمی است که به هر کتاب اختصاص داده می شود و نمی تواند مقدار صفر
دارد یا به و مقداری بین این مقدار تعریف نمی شود.

nominal است. چون روایت یک فرد یا یک کتاب ما را مشخص می کند و مقدار آن
در حد یک کتاب است و عملیات محاسباتی بدون معنی nominal به معنی است و کم
کردن آن به آیدی یک کتاب از نسبی اهمیت خواهد داشت و مقدار معینی هیچ حد نیست
و مقایسه ندارند.

سوال ۳:
(۱) میانگین:

$$\bar{x} = \frac{\sum x}{N} = \frac{10 + 7 + 20 + 12 + 75 + 15 + 9 + 18 + 6 + 12 + 8 + 14}{12} = 17$$

(۲) میانگین استاندارد ها رسمت کرده:

باقیه به اینک تعداد داده ها زوج است میان میانگین دو مقدار وسط است.

$$\text{میان} = \frac{12 + 12}{2} = 12$$

(۳) مد ۱۲ - چون ۱۲ از خردار در بقیه مقدار یکبار رخ داده اند. بین مقدار یک بچین خردار ادا شده ۱۲ است.

(۴) انحراف معیار:

$$s^2 = \frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{12} \left[\frac{10-17}{12} + \frac{7-17}{12} + \frac{20-17}{12} + \frac{12-17}{12} + \frac{75-17}{12} + \frac{15-17}{12} + \frac{9-17}{12} + \frac{18-17}{12} + \frac{6-17}{12} + \frac{12-17}{12} + \frac{8-17}{12} + \frac{14-17}{12} \right]$$

انحراف معیار = $\sqrt{325} \approx 18.03$

$s^2 = 325$

Z-Score = $\frac{x - \bar{x}}{s}$

$\frac{10 - 17}{18.03} \approx -0.38$, $\frac{7 - 17}{18.03} \approx -0.55$, $\frac{20 - 17}{18.03} \approx 0.14$

$\frac{12 - 17}{18.03} \approx -0.27$, $\frac{75 - 17}{18.03} \approx 3.21$, $\frac{15 - 17}{18.03} \approx -0.11$, $\frac{9 - 17}{18.03} \approx -0.44$

$\frac{18 - 17}{18.03} \approx 0.05$, $\frac{6 - 17}{18.03} \approx -0.61$, $\frac{12 - 17}{18.03} \approx -0.27$, $\frac{8 - 17}{18.03} \approx -0.49$

$\frac{14 - 17}{18.03} \approx -0.14$

سوال ٤ :
 فاصله اقلیدسی :

$$d(x, y) = \sqrt{(x_0 - y_0)^2 + (x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2} = \sqrt{45} = 3\sqrt{5}$$

فاصله مانیوتون :

$$(x_0 - y_0) + (x_1 - y_1) + (x_2 - y_2) + (x_3 - y_3) = 11$$

Minkowski norm :

$$\sqrt[p]{(x_0 - y_0)^p + (x_1 - y_1)^p + (x_2 - y_2)^p + (x_3 - y_3)^p} = \sqrt[p]{234} \approx 4,2$$

supremum norm :

$$x_0 - y_0 = 2, \quad x_1 - y_1 = 1, \quad x_2 - y_2 = 4, \quad x_3 - y_3 = 2$$

$$\max(2, 1, 4, 2) = 4$$