

بخش 1 Machine Learning :

- A. Supervised Learning و Unsupervised Learning چه تفاوتی دارند؟
- B. چرا Feature Scaling در الگوریتم‌های Machine Learning ضروری است؟
- C. Standardization و Normalization چه تفاوتی دارند؟
- D. چرا Min-Max Normalization برای مقیاس‌بندی داده‌ها استفاده می‌شود؟
- E. Z-Score Normalization چیست و چرا کاربرد دارد؟
- F. Regularization در الگوریتم‌های Machine Learning چیست؟
- G. Overfitting و Underfitting چه مشکلاتی را در Model-building به وجود می‌آورند؟
- H. Cross-Validation چرا در Train/Test Split کاربرد دارد؟
- I. Gradient Descent چگونه کار می‌کند؟
- J. چرا Deep Learning برای پیچیده‌ترین مسائل استفاده می‌شود؟

مدرس: سجاد احمدزاده

ترم بهمن ۱۴۰۳

Day... Month... Year...

Subject...

مباحث ویژه

Machine Learning 1

A. Supervised Learning و Unsupervised Learning چه تفاوتی دارند؟

در نوع اصلی از یادگیری ماشین هستند که تفاوت‌های اساسی باهم دارند

در روش Supervised Learning مدل با استفاده از داده‌های برچسب دار آموزش می‌بیند

یعنی هر داده ورودی دارای یک خروجی مشخص است و مدل سعی می‌کند رابطه بین

ورودی و خروجی را یاد بگیرد. مثال: تشخیص ایمیل‌های اسپم (مدل با ایمیل‌هایی

که برچسب "اسپم" یا "غیر اسپم" دارند آموزش داده می‌شوند.)

در روش Unsupervised Learning مدل روی داده‌های بدون برچسب آموزش

داده می‌شود و هدف آن کشف الگوها و ساختارهای مخفی در داده‌ها است. مثال:

خوشه بندی مشتریان یک فروشگاه بر اساس الگوهای خرید (مدل بدون دانش برچسب مشخص

مشتریان مشابه را در گروه‌های جداگانه قرار می‌دهد) و

B. چرا Feature Scaling در الگوریتم‌های Machine Learning ضروری است؟

AVANGE

Day... Month... Year...

Subject...

1. Feature scaling فراگیری است که در آن مقادیر ویژگی‌های ورودی در یک بازه مشخص

2. 3. نورمال سازی یا استاندارد سازی می‌شوند. این کار به دلایل زیر در الگوریتم‌های یادگیری ماشین

4. 5. اهمیت زیادی دارد. 1- جلوگیری از تأثیر مقیاس‌های مختلف ویژگی‌ها

6. 7. 2- بهبود عملکرد الگوریتم‌های مبتنی بر فاصله 3- تسریع در همگرایی الگوریتم‌های مبتنی

8. 9. بر پارامترها 4- بهبود پایداری و عملکرد PCA

10. 11. در یک روش‌های رایج برای این کار Normalization و Standardization

12. 13. هستند که بسته به خروج داده و الگوریتم انتخاب می‌شوند.

14. 15. C. Standardization و Normalization چه تفاوتی دارند؟

16. 17. استاندارد سازی مقادیر را طوری تغییر می‌دهد که میانگین داده‌ها صفر و انحراف معیار

18. 19. برابر با یک شود. فرمول: $\frac{x - \mu}{\sigma}$

20. 21. که در آن μ = میانگین ویژگی و σ = انحراف معیار

22. 23. در نورمال سازی داده‌ها را در یک بازه مشخص (معمولاً بین -1 و 1) قرار می‌دهد.

24. AVANCE

Day...	Month...	Year...	Subject...
1			موضوع 1: $\frac{\min X - \max X}{2}$ در آن $\min X$ کمترین مقدار و $\max X$ بیشترین مقدار ویژگی
2			
3			
4			چیز زمانی از کدام روش استفاده کنیم؟
5			
6			از استاندارد سازی استفاده کنیم؟
7			
8			1- داده ها توزیع نرمال دارند 2- در الوریتم های مبتنی بر اربابان استفاده می شود
9			
10			3- داده ها دارای مقدارهای پرت (outliers) هستند برآ استاندارد سازی نیست به
11			
12			مقدار پرت حساس است
13			
14			از نرمال سازی استفاده کنیم؟
15			
16			1- داده ها دارای مقدارهای ^{محدوده} مشخص هستند 2- داده ها توزیع نرمال ندارند و مقدارهای پرت
17			
18			زیادی ندارند 3- الوریتم های مبتنی بر فاصله استاندارد می شوند (مثل KNN و K-Means)
19			
20			D. چرا Min-Max Normalization برای مقیاس بندی داده ها استفاده می شود؟
21			
22			1- مقیاس بندی داده ها به یک محدوده مشخص 2- بهبود عملکرد الوریتم ^{الوریتم} های مبتنی بر فاصله
23			
24			3- بهبود سرعت محاسبه الوریتم های مبتنی بر اربابان 4- مناسب برای داده هایی با محدوده مشخص

AVANCE

Day... Month... Year...

Subject...

1. کاربرد در نمایش داده‌ها و سنجش کلاس‌ها

2. در کل
3. Min-Max Normalization داده‌ها را در یک بازه مشخصی مقیاس بندی می‌کند4.
5. و باعث بهبود عملکرد الگوریتم‌های یادگیری ماشین می‌شود. برای الگوریتم‌های مبتنی بر فاصله6.
7. مانند KNN و K-means و سبک‌های عمومی‌تری است.8.
9. Z-score Normalization چیست و چرا کاربرد دارد؟10.
11. Z-score Normalization یکی از روش‌های مقیاس بندی داده‌ها در یادگیری ماشین است12.
13. که داده‌ها را طوری تبدیل می‌کند که میانگین آنها صفر و انحراف معیارشان برابر با یک باشد14.
15. این روش باعث می‌شود که ویژگی‌هایی با مقیاس‌های مختلف تأثیر یکسانی در مدل داشته باشند.16.
17. کاربرد: ۱. مقایسه ویژگی‌هایی با مقیاس‌های مختلف ۲. بهبود عملکرد18.
19. الگوریتم‌های یادگیری ماشین ۳. کاهش تأثیر مقادیر پرت (outliers)20.
21. F. Regularization در الگوریتم‌های Machine Learning چیست؟22.
23. یکی از تکنیک‌های مهم در یادگیری ماشین است که برای جلوگیری از overfitting به کار می‌رود

Day... Month... Year...

Subject...

1. $overfitting$ زمانی رخ می دهد که مدل بیش از حد به داده های آموزشی (توسیم) وابسته شود.
- 2.
3. و نتواند روی داده های جدید عملکرد خوبی داشته باشد.
- 4.
5. در مدل های پیچیده یادگیری ماشین، مانند شبکه های عصبی عمیق یا رگرسیون چند متغیره،
- 6.
7. احتمال دارد مدل به فیت های موجود در داده ها حساس شود و الگوهای غیر ضروری را یاد بگیرد.
- 8.
9. این موضوع باعث می شود که مدل روی داده های آموزشی عملکرد خوبی داشته باشد اما روی داده های
- 10.
11. جدید ضعیف عمل کند. $Regularization$ یکی کف تا مدل ساده تر شده و تعمیم پذیری
- 12.
13. بهتری داشته باشد.
- 14.
15. روش های $Regularization$:
- 16.
17. 1- $L1 Regularization (Lasso Regression)$
- 18.
19. 2- $L2 Regularization (Ridge Regression)$
- 20.
21. G . $overfitting$ و $underfitting$ به مشکلاتی در $Model-building$ مربوط می آید؟
- 22.
23. $overfitting$ و $underfitting$ دو مشکل رایج در مدل سازی یادگیری ماشین هستند که
- 24.

AVANGE

Day... Month... Year...

Subject...

- 1 می تواند عملکرد مدل را باعث کاهش دهنده
- 2
- 3 *overfitting* : 1- مدل بیش از حد به داده های آموزش وابسته می شود و الگوهای خاص
- 4
- 5 آنرا حفظ می کند 2- در داده های جدید عملکرد ضعیفی دارد چون به نویز و جزئیات
- 6
- 7 غیر ضروری حساس شده است 3- باعث کاهش قابلیت تعمیم مدل می شود
- 8
- 9 *underfitting* : مدل به اندازه کافی پیچیدگی ندارد و نمی تواند الگوهای موجود در داده
- 10
- 11 را بگیرد 2- هم در داده های آموزش و هم در داده های تست عملکرد ضعیفی دارد
- 12
- 13 راه حل ها :
- 14
- 15 برای جلوگیری از *overfitting* به کاهش پیچیدگی افزایش داده ها استفاده از
- 16 *Regularization (L1/L2)*
- 17
- 18 *Cross-Validation* چرا در *Train/Test split* کاربرد دارد
- 19
- 20 1- ارزیابی بهتر مدل 2- جلوگیری از وابستگی به یک تقسیم خاص 3- استفاده
- 21
- 22 موثرتر از داده ها 4- کمک به تنظیم های پارامترها
- 23
- 24 در مدل *Cross-Validation* باعث می شود *Train/Test split* کارکردش را

AVANGE

Day... Month... Year...

Subject...

1 و مدل به جای یکبار تست شدن، روی چندین مجموعه ارزیابی شود که به بهبود تخمین پذیری

3 کمک می کند.

5 Gradient Descent. I چگونه کار می کند؟

7 Gradient Descent به الگوریتم بهینه سازی است که برای کمینه کردن Loss function

9 در مدل های یادگیری ماشین و شبکه های عصبی استفاده می شود. این روش به روزرسانی مداوم

11 وزن ها را انجام می دهد تا مدل بهترین عملکرد را داشته باشد.

13 مراحل اجرای Gradient Descent: 1- مقداردهی اولیه 2- محاسبه گرادیان

15 3- روزرسانی پارامترها 4- تکرار

17 انواع Gradient Descent:

19 Batch Gradient Descent به از یک داده های آموزش برای هر روزرسانی استفاده می کند. (کمتر ولی پایدار)

21 Stochastic Gradient Descent به از یک نمونه تصادفی برای هر روزرسانی استفاده می کند. (سریع ولی ناپایدار)

23 Mini-Batch Gradient Descent به از یک زیرمجموعه کوچک از داده ها در هر مرحله استفاده می کند.

AVANCE

Day... Month... Year...

Subject...

1. چرا *Deep Learning* برای پیچیده ترین مسائل استفاده می شود؟
- 2.
3. *Deep Learning* یکی از قوی ترین روش های یادگیری ماشین است که برای حل مسائل
- 4.
5. بسیار پیچیده مانند پردازش تصاویر، پردازش زبان طبیعی (NLP)، بازی های هوش مصنوعی
- 6.
7. و تشخیص های مالی استفاده می شود. دلیل قدرت یادگیری عمیق این است که می تواند
- 8.
9. ویژگی های سطح بالا را از داده ها استخراج کند.
- 10.
11. دلایل استفاده از *Deep Learning* در مسائل پیچیده:
- 12.
13. ۱- توانایی یادگیری ویژگی ها به صورت خودکار ^{قابلیت} ۲- پردازش حجم عظیمی از داده ها
- 14.
15. ۳- مدل سازی روابط غیر خطی پیچیده ۴- یادگیری سیستم مراقبتی از داده ها
- 16.
17. ۵- موفقیت در کاربردهای پیشرفته
- 18.
19. در *Deep Learning* به دلیل قدرت یادگیری ویژگی ها، پردازش داده های
- 20.
21. عظیم، وصول سازی روابط پیچیده، برای پیچیده ترین مسائل هوش مصنوعی
- 22.
23. استفاده می شود.
- 24.