

جزوه دوره (نسخه پژوهشی)

جامپ یادگیری ماشین (Quera College)

نویسنده: نام شما

تاریخ: ۱۴۰۴ اسفند

فهرست مطالب

۶

پیشگفتار

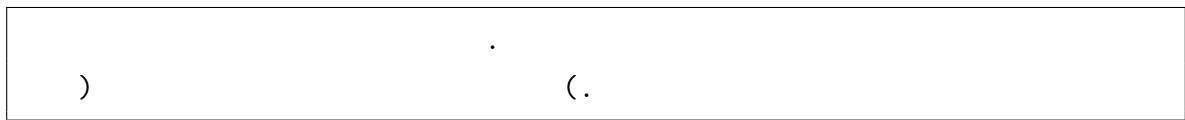
۷	۴ مهندسی ویژگی	
۷	۱.۴	اهمیت
۷	۲.۴	مشخصات ویژگی خوب
۷	۳.۴	ویژگی‌های دسته‌ای
۷	۴.۴	مقادیر گم شده در ویژگی‌های دسته‌ای
۷	۵.۴	ویژگی‌های تقویمی
۸	۶.۴	سنتر ویژگی
۸	۷.۴	تغییر مقیاس ویژگی
۸	۸.۴	نشت داده
۸	۹.۴	فوت و فن‌های مهندسی ویژگی
۸	۱۰.۴	کاهش ابعاد
۸	۱۱.۴	انتخاب ویژگی
۸	۱۲.۴	خط لوله
۹	۵ رگرسیون	
۹	۱.۵	اهداف فصل
۹	۲.۵	مقدمه
۹	۳.۵	مدل چیست؟
۹	۴.۵	تخمین، تابع هزینه و بهینه‌سازی
۹	۵.۵	رگرسیون خطی
۱۰	۶.۵	ارزیابی
۱۰	۷.۵	رگرسیون چندجمله‌ای
۱۰	۸.۵	عمومیت
۱۰	۹.۵	رگولاریزیشن
۱۱	۶ دسته‌بندی	
۱۱	۱.۶	مقدمه
۱۱	۲.۶	رگرسیون لجستیک
۱۱	۳.۶	ارزیابی - قسمت اول
۱۱	۴.۶	ارزیابی - قسمت دوم
۱۱	۵.۶	کراس ولیدیشن
۱۲	۶.۶	نزدیکترین- k همسایه
۱۲	۷.۶	بیز ساده‌لوحانه
۱۲	۸.۶	ماشین بردار پشتیبان
۱۲	۹.۶	هایبری‌پارامترها
۱۲	۱۰.۶	آشنایی با کتابخانه‌ی OVR
۱۲	۱۱.۶	درخت تصمیم
۱۲	۱۲.۶	فوت و فن درخت تصمیم
۱۳	۱۳.۶	بیش‌برازش درخت تصمیم

<p>۱۴</p> <p>۱۴</p> <p>۱۴</p> <p>۱۴</p> <p>۱۴</p> <p>۱۴</p> <p>۱۵</p> <p>۱۵</p>	<p>۷ یادگیری تجمعی</p> <p>۱.۷ اهداف فصل</p> <p>۲.۷ مقدمه</p> <p>۳.۷ جنگل تصادفی</p> <p>۴.۷ الگوریتم AdaBoost</p> <p>۵.۷ الگوریتم Boosting Gradient</p> <p>۶.۷ الگوریتم XGboost</p> <p>۷.۷ روش Stacking</p>
<p>۱۶</p> <p>۱۶</p> <p>۱۶</p>	<p>۸ پروژه اول</p> <p>۱.۸ مقدمه</p> <p>۲.۸ یادداشت‌ها و راه حل</p>
<p>۱۷</p> <p>۱۷</p> <p>۱۷</p> <p>۱۷</p> <p>۱۷</p> <p>۱۷</p>	<p>۹ شبکه عصبی</p> <p>۱.۹ اهداف فصل</p> <p>۲.۹ پرسپترون</p> <p>۳.۹ آموزش پرسپترون</p> <p>۴.۹ پرسپترون چندلایه</p> <p>۵.۹ عمومیت</p>
<p>۱۸</p> <p>۱۸</p> <p>۱۸</p> <p>۱۸</p> <p>۱۸</p> <p>۱۸</p> <p>۱۹</p>	<p>۱۰ یادگیری ناظارت‌نشده</p> <p>۱.۱۰ مقدمه</p> <p>۲.۱۰ الگوریتم PCA</p> <p>۳.۱۰ الگوریتم t-SNE</p> <p>۴.۱۰ خوشه‌بندی با k-means</p> <p>۵.۱۰ خوشه‌بندی با k-modes</p> <p>۶.۱۰ خوشه‌بندی با k-prototype</p>
<p>۲۰</p> <p>۲۰</p> <p>۲۰</p> <p>۲۰</p> <p>۲۰</p>	<p>۱۱ پروژه دوم</p> <p>۱.۱۱ اهداف فصل</p> <p>۲.۱۱ تعبیه‌ی متن</p> <p>۳.۱۱ فاصله‌ی ویرایش</p> <p>۴.۱۱ معیار شباهت RBO</p>
<p>۲۱</p> <p>۲۱</p> <p>۲۱</p> <p>۲۱</p>	<p>۱۲ بیشتر بدانید</p> <p>۱.۱۲ نمونه‌کاهی با NearMiss</p> <p>۲.۱۲ نمونه‌افزایی با SMOTE</p> <p>۳.۱۲ درخت رگرسیون</p>
<p>۲۲</p>	<p>واژه‌نامه</p>

فهرست تصاویر

فهرست جداول

پیشگفتار



فصل ۱

مقدمه

۱.۱ سلام!

سلام؛ ورود شما را به کالج «یادگیری ماشین ۲ | جامپ تکنیکال» خوش آمد می‌گوییم. از اینکه در کوئرا کالج افتخار میزبانی شما را داریم، به خود می‌باليم

«یادگیری ماشین ۲ | جامپ تکنیکال» کالج سوم از مسیر علم داده و یادگیری ماشین کوئراست که پس از «یادگیری ماشین ۰ | دروازه ورود» و «یادگیری ماشین ۱ | تحلیل داده با پایتون» طراحی شده است.

هدف ما از تدوین کالج این است که شما را به شکل اصولی و گامبهگام با الگوریتم‌های یادگیری ماشین آشنا کنیم؛ به طوری که در نهایت بتوانید الگوریتم‌ها را تحلیل کنید، نقاط ضعف و قوت آن‌ها را بشناسید و تشخیص دهید چگونه از آن‌ها برای حل مسائل دنیای واقعی کمک بگیرید! علاوه بر این‌که با الگوریتم‌های یادگیری ماشین می‌شوید، تکنیک‌هایی برای استفاده عملی از آن‌ها را نیز خواهید آموخت؛ به عبارت بهتر، «یادگیری ماشین ۲ | جامپ تکنیکال» آمیخته‌ای موزون از دانش علمی و مهارت عملی یادگیری ماشین کلاسیک است!

این کالج مناسب افرادی است که در حد متوسط با پایتون و کتابخانه‌های Numpy و Pandas آشنایی داشته باشند. داشتن دانش حداقلی از آمار احتمال و حسابان (مباحث مشتق) و جبر خطی (ماتریس‌ها) به شما در فرآگیری محتوای این کالج کمک می‌کند. به صورت کلی می‌توان گفت اگر با ریاضیات در حد مقطع دبیرستان آشنا باشید، به راحتی می‌توانید به ماجراجویی در این کالج بپردازید! اگر با پیشنبازهای کالج آشنا نیستید، می‌توانید با گذراندن سایر دوره‌های کوئرا کالج، مهارت‌های لازم را کسب کنید.

شما با داشتن پیش‌نبازهای کالج شروع می‌کنید و در انتهای پس از مطالعه‌ی درسنامه‌ها و حل تمرین‌ها، به «دانشمند داده» و «مهندس یادگیری ماشین» تبدیل می‌شوید که قطعاً راه طولانی اما شیرینی برای حرفة‌ای شدن پیش رو دارید!

البته برای حفظ تناسب کالج و افزایش احتمال یادگیری عميق، از آموزش کامل یادگیری عميق، که زیرمجموعه یادگیری ماشین است، پرهیز شده است؛ بلکه شبکه‌های عصبی مصنوعی و مقدمات یادگیری عميق را در فصل «شبکه عصبی» آورده‌ایم. برای فرآگیری یادگیری عميق، می‌توانید منتظر کالج‌های بعدی مسیر علم داده و یادگیری ماشین کوئرا کالج باشید

جزوهٔ جامپ یادگیری ماشین

امیدواریم با ارائهٔ آموزش باکیفیت، گامی مثبت در افزایش دانش و پیشرفت شما برداریم. سختی‌های مسیر نه تنها ناامیدمان نمی‌کند بلکه توانمان را بیشتر و تصمیم‌مان را راسخ‌تر می‌کند. مسیر رسیدن به قله‌ای که آرزویش را داریم؛ جوانانی توانمند، پرتشاش و ایرانی پیشرفت...



۲.۱ قالب کالج

[]

۳.۱ سیستم امتیازات و دریافت گواهی اصلی

[]

۴.۱ یادگیری ماشین چیست؟

[]

۵.۱ چرا پایتون؟

]

[

۶.۱ آماده‌سازی محیط کار

]

[

۷.۱ اجرای نوت‌بوک‌ها

]

[

۸.۱ گوگل کولب

]

[

۹.۱ معرفی مجموعه‌داده

]

[

فصل ۲

مدیریت پروژه

۱.۲ اهداف فصل

] [

۲.۲ چرخه پروژه

] [

۳.۲ اولویت‌بندی

] [

۴.۲ سازماندهی تیم

] [

۵.۲ چرا پروژه‌ها شکست می‌خورند؟

] [

فصل ۳

آماده‌سازی داده

۱.۳ اهداف فصل

[]

۲.۳ سوالاتی درباره داده

[]

۳.۳ چالش‌های داده

[]

۴.۳ ویژگی‌های مجموعه‌داده‌ی خوب

[]

۵.۳ تقسیم‌بندی مجموعه‌داده

[]

۶.۳ داده‌های پرت

] [

۷.۳ مقادیر گم شده

] [

۸.۳ مجموعه داده نامتوازن

] [

فصل ۴

مهندسی ویژگی

۱.۴ اهمیت

[]

۲.۴ مشخصات ویژگی خوب

[]

۳.۴ ویژگی‌های دسته‌ای

[]

۴.۴ مقادیر گم‌شده در ویژگی‌های دسته‌ای

[]

۵.۴ ویژگی‌های تقویمی

[]

۶.۴ سنتز ویژگی

] [

۷.۴ تغییر مقیاس ویژگی

] [

۸.۴ نشت داده

] [

۹.۴ فوت و فن‌های مهندسی ویژگی

] [

۱۰.۴ کاهش ابعاد

] [

۱۱.۴ انتخاب ویژگی

] [

۱۲.۴ خط لوله

] [

فصل ۵

رگرسیون

۱.۵ اهداف فصل

] [

۲.۵ مقدمه

] [

۳.۵ مدل چیست؟

] [

۴.۵ تخمین، تابع هزینه و بهینه‌سازی

] [

۵.۵ رگرسیون خطی

] [

۶.۵ ارزیابی

[]

۷.۵ رگرسیون چندجمله‌ای

[]

۸.۵ عمومیت

[]

۹.۵ رگولاریزیشن

[]

فصل ۶

دسته‌بندی

۱.۶ مقدمه

]

]

۲.۶ رگرسیون لجستیک

]

۳.۶ ارزیابی - قسمت اول

]

۴.۶ ارزیابی - قسمت دوم

]

۵.۶ کراس ولیدیشن

]

۶.۶ نزدیکترین- k -همسايه

] [

۷.۶ بيز ساده‌لوحانه

] [

۸.۶ ماشين بردار پشتيبان

] [

۹.۶ هايپرپaramترها

] [

۱۰.۶ آشنائي با کتابخانه‌ي O2H

] [

۱۱.۶ درخت تصميم

] [

۱۲.۶ فوت و فن درخت تصميم

] [

۱۳.۶ بیشبرازش درخت تصمیم

]

فصل ٧

يادگیری تجمعی

١.٧ اهداف فصل

] [

٢.٧ مقدمه

] [

٣.٧ جنگل تصادفی

] [

٤.٧ الگوریتم AdaBoost

] [

٥.٧ الگوریتم Boosting Gradient

] [

۶.۷ الگوریتم XGboost

]

[

۷.۷ روش Stacking

]

[

فصل ۸

پروژه اول

۱.۸ مقدمه

] [

۲.۸ یادداشت‌ها و راه حل

] [

فصل ۹

شبکه عصبی

۱.۹ اهداف فصل

]
[

۲.۹ پرسپترون

]
[

۳.۹ آموزش پرسپترون

]
[

۴.۹ پرسپترون چندلایه

]
[

۵.۹ عمومیت

]
[

فصل ۱۰

یادگیری نظارت‌نشده

۱.۱۰ مقدمه

]

۲.۱۰ الگوریتم PCA

]

۳.۱۰ الگوریتم t-SNE

]

۴.۱۰ خوشبندی با k-means

]

۵.۱۰ خوشبندی با k-modes

]

۶.۱۰ خوشبندی با k-prototype

]

فصل ۱۱

پروژه دوم

۱.۱۱ اهداف فصل

]
[

۲.۱۱ تعبیهی متن

]
[

۳.۱۱ فاصلهی ویرایش

]
[

۴.۱۱ RBO معیار شباهت

]
[

فصل ۱۲

بیشتر بدانید

۱.۱۲ نمونه‌کاهی با NearMiss

]

]

۲.۱۲ نمونه‌افزایی با SMOTE

]

۳.۱۲ درخت رگرسیون

]

واژه‌نامه

() .

منابع

() . / / /