محتوای دوره عمومی دیتاساینس:

جلسه‌ی اول:

مقدمه و پیش‌نیازها

جلسه‌ی دوم:

بخش اول، احتمالات

بخش دوم، ترکیب‌شناسی

بخش سوم، پایتون – متغیرها و معرفی برخی از توابع

توضیحات تکمیلی

جلسه‌ی سوم:

بخش اول، احتمالات – قانون بیز

بخش دوم، پایتون – عملگرها

جلسه‌ی چهارم:

بخش اول، توزیع‌های احتمال

بخش دوم، توزیع‌های گسسته

بخش سوم، توزیع‌های پیوسته

بخش چهارم، پایتون – شرطی‌ها

جلسه‌ی پنجم:

توابع در پایتون

جلسه‌ی ششم:

بخش اول، یک مثال عملی، کاربرد احتمالات

بخش دوم، سری‌ها در پایتون

جلسه‌ی هفتم:

بخش اول، چرا احتمالات

بخش دوم، حلقه‌های تکرار در پایتون

بخش سوم، مباحث پیشرفته در پایتون‌

جلسه‌ی هشتم:

آمار توصیفی و تحلیل داده‌ها

دنياي ماركتينگ (اختیاری)

یک مثال کاربردی از آمار توصیفی در تحلیل داده‌ها

جلسه‌ی نهم:

آمار استنتاجی در علم داده

جلسه‌ی دهم:

بازه اطمینان – بخش اول

بازه اطمینان بخش دوم

بازه اطمینان – بخش سوم

یک مثال کاربردی از بازه اطمینان در دنیای مارکتینگ

جلسه‌ی یازدهم:

آزمون فرض در ديتاساينس

خطاها و p-مقدار

آزمون فرض براي چند جامعه

تشخیص شکاف حقوقی بین کارمندان زن و مرد یک کمپانی با استفاده از آزمون فرض

جلسه‌ی دوازدهم:

رگرسیون خطی ساده و پیاده‌سازی آن در پایتون، بخش اول، تئوری

رگرسیون خطی ساده و پیاده‌سازی آن در پایتون، بخش دوم، کدنویسی

جلسه‌ی سیزدهم:

رگرسیون چندگانه

ضریب تعیین و کاربردش

پیاده‌سازی رگرسیون خطی چندگانه در پایتون

ضریب تعیین تعدیل شده و آزمون اف

مفروضات رگرسیون

جلسه‌ی چهاردهم:

رگرسیون چندگانه و داده‌های مجازی

پیش‌بینی با استفاده از مدل رگرسیون چندگانه

رگرسیون چندگانه و مدل پیش‌بینی خانه

جلسه‌ی پانزدهم:

ورود به دنیای سای‌ کیت لرن

رگرسیون خطی ساده با سای کیت لرن

رگرسیون خطی چندگانه با سای کیت لرن

پیش‌بینی رگرسیون چندگانه با داده‌های استاندارد

جلسه‌ی شانزدهم:

جداسازی بصورت آموزش و تست

پروژه‌ی پیش‌بینی یا برآورد قیمت ماشین‌های دست‌دوم (پیش پردازش)

جلسه‌ی هفدهم:

پروژه‌ی پیش‌بینی یا برآورد قیمت ماشین‌های دست‌دوم (بررسی مفروضات رگرسیون)

پروژه‌ی پیش‌بینی قیمت ماشین (افزودن متغیرهای مجازی (Dummy))

جلسه‌ی هجدهم:

پروژه‌ی پیش‌بینی قیمت ماشین (معرفی مدل رگرسیون)

پروژه‌ی پیش‌بینی قیمت ماشین (آزمودن مدل رگرسیون)

مشاهده‌ی یک مدل پیاده‌شده روی یک سایت ایرانی

جلسه‌ی نوزدهم:

ورود به رگرسیون لجستیک

مدل رگرسیون لجستیک و لاجیت

جلسه‌ی بیستم:

بررسی برخی از عناصر جدول مدل رگرسیون

درستی و تست مدل رگرسیون لجستیک

جلسه‌ی بیست و یکم:

مثالی از رگرسیون لجستیک با سای کیت لرن (scikit-learn)

ورود به دنیای خوشه‌بندی

پروژه‌ی علم داده – یافتن مدلی برای قبولی دانش‌آموزان

جلسه‌ی بیست و دوم:

خوشه‌بندی به روش K-Means

خوشه‌بندی به روش K-Means در محیط ژوپیتر

جلسه‌ی بیست و سوم:

نکاتی در خصوص خوشه‌بندی و استانداردسازی ویژگی‌ها

یافتن بهینه‌ترین تعداد از خوشه‌ها با استفاده از WCSS

جداسازی بازار (Market Segmentation) و خوشه‌بندی در محیط ژوپیتر

پروژه‌ی علم داده – بکارگیری خوشه‌بندی در طبقه‌بندی کشورها و دیتاست گیاه زنبق (iris dataset)

جلسه‌ی بیست و چهارم:

خوشه‌بندی سلسه مراتبی و نمودار درختی

مثالی از خوشه‌بندی سلسله مراتبی و نمودار درختی در محیط ژوپیتر

جلسه‌ی بیست و پنجم:

یادگیری ماشین و عمیق به زبان ساده

جلسه‌ی بیست و ششم:

اجزای یادگیری ماشین

یک مثال ملموس از کارکرد نرخ یادگیری (در محیط اکسل)

ساختن یک الگوریتم یادگیری ماشین از صفر با نامپای

جلسه‌ی بیست و هفتم:

نکته‌ی تکمیلی در خصوص ساختن الگوریتم با نامپای

نکاتی در خصوص تنسورفلو و طریقه‌ی نصب آن

نوشتن یک الگوریتم یادگیری ماشین با استفاده از تنسورفلو

جلسه‌ی بیست و هشتم:

مقدمه‌ای بر شبکه‌های عصبی عمیق (همراه با معرفی کتاب)

جلسه‌ی بیست و نهم:

بیش‌برازش و روش‌های شناخت و رفع آن

جلسه‌ی سی‌ام:

مقداردهی اولیه به وزن‌ها (Initialization)

گرادیان کاهشی تصادفی و تکانه (SGD and momentum)

نکاتی تکمیلی در خصوص نرخ یادگیری

آشنایی با سه بهینه‌ساز بسیار مهم در یادگیری ماشین

جلسه‌ی سی‌ و یکم:

نکاتی تکمیلی در خصوص پیش‌پردازش

جلسه‌ی سی‌ و دوم:

نکاتی در خصوص دیتاست MNIST

پروژه‌ی استخراج اعداد از دیتاست MNIST با استفاده از تنسورفلو

جلسه‌ی سی‌ و سوم (آخرین جلسه‌ی دوره):

پروژه‌ی پیش‌بینی خرید مجدد مشتریان از یک پلت‌فرم فروش کتاب صوتی (آشنایی با داده‌ها)

پروژه‌ی پیش‌بینی خرید مجدد مشتریان از یک پلت‌فرم فروش کتاب صوتی (کدهای پروژه در تنسورفلو)