

سرفصل های دوره:

۱- معرفی کلی زبان برنامه نویسی و ابزار های آن

▪ آموزش برنامه نویسی، آشنایی و معرفی مبانی پایتان

▪ معرفی منو های مختلف، ابزار کاربردی و مهم پایتان

۲- بررسی توابع و دستور ها، متغیر ها، عملگر ها، شرط ها و حلقه ها، رشته ها و ساختمان داده

▪ نوشتن تابع در پایتان و بررسی انواع توابع

▪ عملگر های منطقی، مقایسه ای، بیتی و شرطی و اولویت آن ها

▪ حلقه های تو در تو و ساختمان داده

▪ ساختار های شرطی If, Else

▪ حلقه های تکرار For, While

▪ ایجاد لیست های مختلف، جست و جو و حذف در لیست ها، بررسی اسلایس ها

▪ نحوه تعریف رشته و تجزیه و تحلیل متد های پایتان

▪ دیکشنری ها، ماتریس ها و متد ها در دیکشنری ها، معرفی نوع داده در دیکشنری ها

▪ ورودی اطلاعات از منابع مختلف همچون اکسل، متلب

▪ بررسی انواع خطاها در پایتان ، فرمت بندی خروجی ها

۳- استفاده از کتابخانه ها و پکیج های علم داده، یادگیری ماشین و داده کاوی در Python

▪ بررسی پکیج Numpy به منظور پیاده سازی عملیات ریاضی

▪ بررسی پکیج Scipy به منظور پیاده سازی عملیات جبری و آماری

▪ بررسی پکیج Pandas به منظور تحلیل دیتافریم ها و ورودی داده ها به پایتان

▪ بررسی پکیج یادگیری ماشین و داده کاوی Sklearn و Mlxtend

▪ بررسی پکیج های تحلیل اکتشافی داده نظیر Pandasgui و Pandas_Profiling

▪ بررسی پکیج مصورسازی داده نظیر Matplotlib و Seaborn

▪ بررسی پکیج Plotly بمنظور داستان سرایی و داشبوردسازی داده

۴- مصورسازی داده و رسم انواع نمودارهای دو بعدی و سه بعدی و ماژول های گرافیکی

▪ نحوه ایجاد یک ماژول

▪ انواع ماژول، استفاده از تابع در ماژول، جست و جو در یک ماژول

▪ رسم نمودار های ۲ بعدی، هیستوگرام، پله ای، میله ای، خطی، ناحیه ای، نمودار پای چارت، نمودار پراکندگی ی، نمودار های حبابی،

۵- نمودار جعبه ای و شناور، نمودار ویالونی، نمودار نقشه جهان

▪ رسم نمودار های ۳ بعدی

▪ ترسیم نمودار ها بصورت مجزا و کنار هم

۶- علم داده، داده کاوی و یادگیری ماشین

▪ معرفی و تشریح تفاوت های علم داده، یادگیری ماشینی، داده کاوی، هوش تجاری و عظیم داده ها

▪ جایگاه زبان برنامه نویسی پایتان در حوزه علم داده

▪ نیازها و چالش های موقعیت های شغلی مختلف علم داده با توجه به یادگیری زبان پایتان

▪ مرور اجمالی بر محیط های توسعه IDE و بیان نقاط ضعف و قوت هر کدام

▪ راه اندازی پایتان، معرفی کلی زبان برنامه نویسی پایتان، معرفی انواع کتابخانه ها و چگونگی طریقه ورود داده از منابع مختلف

▪ ارائه مبانی ابتدایی در محیط برنامه نویسی پایتان

▪ نصب و راه اندازی کتابخانه های Scikit-Learn جهت پیاده سازی الگوریتم های یادگیری ماشینی

▪ تفاوت Verification و Validation

▪ تفاوت Data Gathering و Data Collecting تفاوت های Data Base, DataSet, Data Frame

▪ تفاوت Data Analytics و Data Analysis

• تفاوت داده های Train و Test و Validation

• تفاوت مفاهیم Overfitting و Underfitting

۷- داده کاوی و یادگیری باناظر و بدون ناظر در پایتان:

• بررسی و تحلیل مسائل طبقه بندی (Classification) بر روی حوزه های مختلف من جمله سلامت، پتروشیمی و خودروسازی

• پیاده سازی الگوریتم درخت تصمیم و جنگل های تصادفی Decision Tree & Random Forest

• الگوریتم ماشین های بردار پشتیبان Support Vector Machine

• الگوریتم نزدیک ترین همسایگی KNN

• الگوریتم نایو-بیز Naïve Bayes

• الگوریتم شبکه های عصبی Neural Networks

• الگوریتم رگرسیون لجستیک Logistic Regression

• پیاده سازی الگوریتمهای یادگیری تجمیعی (Ensemble Learning) شامل:

Bagging و Boosting

• بررسی و تحلیل مسائل رگرسیون (Regression)

• پیاده سازی الگوریتم های رگرسیونی شامل رگرسیون خطی و چند متغیره

• بررسی و تحلیل مسائل مسائل خوشه بندی (Clustering)

• پیاده سازی الگوریتمهای DBscan و K-Means , K mediods

• بررسی و تحلیل مسائل مسائل قواعد انجمنی (Association rules) با استفاده از تکنیک Apriori در پایتان

• ارزیابی مدل - امتیاز دهی مدل و اعتبار سنجی (K Fold Cross Validation)

• پیاده سازی روش های اعتبار سنجی و بهبود مدل

• بهینه سازی الگوریتم ها و توسعه مدل