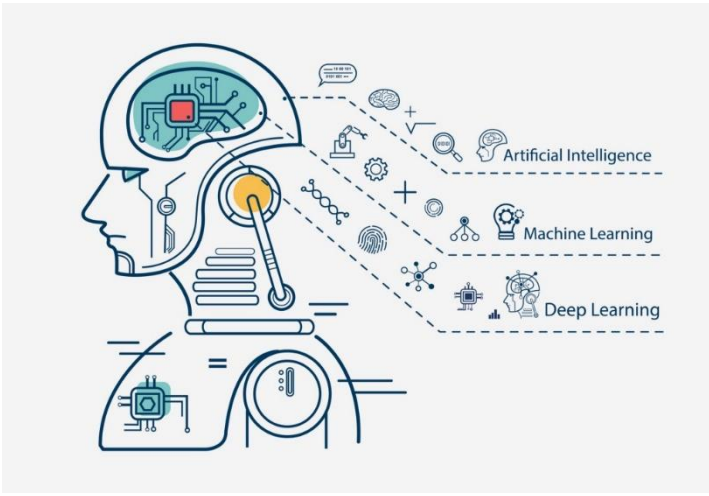


شناسنامه دوره آموزشی بیگ دیتا و یادگیری ماشینی



نام دوره: بیگ دیتا و یادگیری ماشینی

سطح دوره: تخصصی

مخاطبین دوره: کارشناسان داده

نوع دوره: کارگاهی

مدت دوره: ۴۰ ساعت

پیش نیاز دوره: بدون پیش نیاز

نحوه برگزاری دوره: ۵ جلسه ۸ ساعته

معرفی دوره:

با گسترش کاربردهای فناوری اطلاعات در حوزه های مختلف، نیاز به خودکارسازی فرآیندهای تصمیم سازی و تصمیم گیری، روند فزاینده ای داشته است. دانش هوش مصنوعی به عنوان یکی از راهکارهای اصلی برای رفع این نیازها، از روشهای مبتنی بر یادگیری ماشینی استفاده میکند.

با استفاده از تکنیک های یادگیری ماشین، کامپیوتر، الگوهای موجود در داده ها (اطلاعات پردازش شده) را یاد گرفته و میتواند از آن استفاده کند. یادگیری ماشین با فرآیند های داده کاوی، بسیار شبیه (و از نگاه کاربردی تقریباً یکسان) است. در فرآیند یادگیری ماشین، داده ها، بسیار اهمیت دارند. اصطلاح معروفی در این حوزه وجود دارد که به این صورت ترجمه کرده ایم: اگر داده ی بد، به سیستم تزریق شود، خروجی نیز، خروجی بدی خواهد بود به این معنی که، هر چقدر الگوریتم های مختلف یادگیری ماشین، قوی و جامع طراحی شوند، اگر داده های خوبی به سیستم وارد نشود (مثلا داده های غلط یا داده های ناکافی)، سیستم، پاسخی غیر دقیق و ناصحیح ارائه می دهد.

شبکه های عصبی عمیق نیز یکی از مباحث مهم و پرکاربرد نوین در علوم کامپیوتر و یادگیری ماشین است. این دسته از شبکه ها که قابلیت یادگیری الگوهای پیچیده را نیز دارند، به خاطر قدرت و دقت بالایشان در بسیاری از مسائل دنیای واقعی (مانند تشخیص پلاک، پیش بینی بیماری ها، دسته بندی خودکار متون و تصاویر و...) به کار گرفته شده اند. امروزه صنعت هوش مصنوعی به شدت در حال گسترش است و سرمایه گذاری های عظیمی در سطح دنیا روی این مبحث در حال انجام گرفتن است. شاخص ترین شاخه در حوزه ی هوش مصنوعی، یادگیری ماشین و یادگیری عمیق است که هدف از آن ایجاد قابلیت یادگیری توسط ماشین ها میباشد به گونه ای که بتوانند همانند انسان و حتی بهتر از انسان ها فکر کنند.

اهداف دوره:

تربیت دانشمند داده و داده کاو به منظور استخراج دانش از داده های موجود سازمان

سر فصل دوره:

بخش اول: مبانی یادگیری ماشینی

- مقدمه ای بر یادگیری ماشینی
- یادگیری با نظارت، دسته بندی با استفاده از الگوریتم KNN، روش های مختلف محاسبه فاصله، درخت تصمیم، مسئله تقریب تابع
- یادگیری بی نظارت، خوشه بندی با استفاده از K-Means، خوشه بندی سلسله مراتبی
- کاهش ابعاد، آشنایی با PCA، آشنایی با SVD
- ماشین بردار پشتیبانی
- نحوه ارزیابی مدل، مفهوم بیش برآزش و زیر برآزش
- معیارهای ارزیابی، دقت، یادآوری، صحت، ROC، ماتریس برخورد

بخش دوم: کلیات و مفاهیم پایه در یادگیری ماشینی

- ۱-۲- تعاریف
 - یادگیری تحت نظارت
 - یادگیری بدون نظارت
 - ۲-۲- دسته بندی (Classification)
 - ۲-۳- خوشه بندی (Clustering)
 - تکنیک های محاسبه فاصله بین انواع ویژگی ها
 - روش خوشه بندی K-Means
 - روش خوشه بندی K-Medoids
 - روش های خوشه بندی سلسله مراتبی (Hierarchical)
 - شاخص های ارزیابی فرآیند خوشه بندی
 - ۲-۴- مرور روش های کلاسیک در یادگیری ماشینی
 - یادگیری مبتنی بر نمونه ها (Learning based-Instance)
 - یادگیری مبتنی بر قواعد (Learning based-Rule)
 - یادگیری مبتنی بر نظریه احتمالات (Bayesian Learning)
 - درخت تصمیم
 - الگوریتم ID3
 - الگوریتم C4.5
 - ۲-۵- ترکیب دسته بندها (Combining Classifiers)
 - روش Bagging
 - روش Boosting
 - روش AdaBoost
 - ۲-۶- یادگیری تقویتی Reinforcement Learning

- معرفی مفاهیم پایه (Agent, Action, Policy,...)
- روش های انتخاب کنش
- روش برنامه ریزی پویا
- روش تقویتی مونت کارلو

بخش سوم: مبانی یادگیری عمیق

- ۱-۳ - مقدمه ای بر یادگیری عمیق و شبکه های عصبی
 - یادگیری عمیق چیست؟
 - کاربردهای یادگیری عمیق
 - شواهد زیستی و اتصالات
 - معماری شبکه عصبی
 - معرفی و بررسی انواع توابع فعال سازی (activation function)
 - معرفی مدل های پرسپترون چندلایه MLP یا multi-layer Perceptron
 - انواع توابع هزینه در شبکه های عصبی Cost Functions یا Loss Function
 - یادگیری و آموزش شبکه های عصبی
- ۲-۳ - اکوسیستم پایتون برای یادگیری عمیق و پیاده سازی شبکه های عصبی با کتابخانه های مربوطه
 - معرفی پایتون و کتابخانه های علمی و یادگیری ماشین پایتون
 - آموزش مقدماتی کتابخانه یادگیری ماشین scikit-learn
 - معرفی و آموزش کتابخانه یادگیری عمیق Keras
 - بررسی عملی تفاوت روش های سطحی یادگیری ماشین و یادگیری ماشین با پیاده سازی
- ۳-۳ - شبکه های عصبی همگشت (کانولوشنال) و بینایی ماشین
 - معرفی و آشنایی با رویکرد آموزش کرنل یا فیلترهای همگشت به جای طراحی فیلتر
 - مروری بر همگشت Valid و Same
 - مشکلات همگشت valid در طراحی شبکه های همگشتی و فرمول اندازه خروجی پس از همگشت
 - فرمول محاسبه مقدار padding برای همگشت same
 - معرفی strided convolution
 - همگشت روی کانال های تصویر (همگشت سه بعدی)
 - چند فیلتر در همگشت و خروجی چند کاناله
 - محاسبه تعداد پارامترها در هر لایه همگشت
 - معرفی واحدهای pooling و ادغام بیشینه
 - بررسی چندین معماری معروف همگشتی
 - انتقال یادگیری (Transfer Learning)
- ۳-۴ - پیاده سازی شبکه های عصبی همگشت با پایتون (Keras) و انتقال یادگیری
 - پیاده سازی شبکه عصبی همگشتی در کتابخانه کراس و آشنایی با لایه های Conv2D و MaxPooling2D
 - آشنایی با متد Flatten برای لایه ماقبل Dense
 - بررسی ابعاد activation هر لایه: حساب دستی و آشنایی با متد summary()
 - بررسی Sigmoid و Softmax برای لایه آخر
- ۳-۵ - پردازش های سریع و کاربردهای پردازش موازی در یادگیری عمیق
 - معرفی مفاهیم پردازش موازی و کاربردهای آن
 - انتخاب یک سیستم مناسب برای پردازش های سریع

- معرفی ابزارهای پردازش سریع برای پلتفرم های مختلف
- مقدمه ای بر مفاهیم همزمانی و صف بندی
- مقایسه همزمانی با موازی سازی
- مقدمه ای بر مفاهیم موازی سازی بر روی CPU، TPU، GPU، CUDA و CuDNN
- ۳-۶- یادگیری عمیق بدون نظارت و خودرمنزنگارها (Autoencoder)
- استخراج ویژگی (روش های دستی و یادگیری عمیق)
- ویژگی خوب چه ویژگی ای است؟
- اتوانکدر چیست؟
- اتوانکدر در شاخه های مختلف (متن، تصویر، صوت و ...)
- معرفی و بررسی انواع اتوانکدرها شامل Variational، Sparse، Denoising و Convolutional Stacked
- پیاده سازی اتوانکدرها با کتابخانه کراس
- ۳-۷- یادگیری عمیق برای داده های ساختاریافته (جدولی) و سیستم های پیشنهاددهنده
- تفاوت داده های ساختار یافته و بدون ساختار
- روش های سنتی یادگیری ماشین برای کار با داده های جدولی
- معرفی روش Entity Embedding برای کار با متغیرهای دسته ای Categorical
- پیاده سازی Entity Embedding با استفاده از کتابخانه fastai
- کاربرد یادگیری عمیق برای سیستم های پیشنهاددهنده و Collaborative Filtering