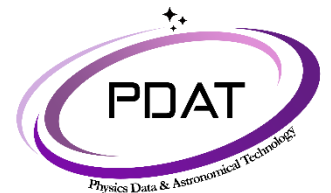




# Data Science

Pouriya Khalilian

Qorpi AI Workgroup  
PDAT Data Science Laboratory



## پوریا فلیلیان

سرگروه تیم هوش مصنوعی کورپی

محقق هوش مصنوعی و سیستم های پیچیده در آزمایشگاه علم داده PDAT



## Publications

- Modeling the Central Supermassive Black Holes Mass of Quasars via LSTM Approach
- Design interpretable experience of dynamical feed forward machine learning model for forecasting NASDAQ
- Modeling using Recurrent Neural Networks with Interpretable Dataset Design for LIBS Spectrum Classification
- Machine learning model to project the impact of Ukraine crisis
- Considering Interpretability of the LSTM Architecture for Oil Stocks Prices Prediction
- Recurrent Neural Networks for continuous prediction of aluminum alloy concentrations in Laser Induced Breakdown Spectroscopy

Research Gate : [researchgate.net/profile/Pouriya-Khaliliyan](https://researchgate.net/profile/Pouriya-Khaliliyan)

Linkedin : [linkedin.com/in/pouriya-khaliliyan](https://linkedin.com/in/pouriya-khaliliyan)

# سرفصل

## ۱- مقدمه علم داده

- آشنایی با علم داده
- علم داده در صنعت و تجارت (هوش تجاری)
- علم داده در تحقیق و ریسرچ
- پروژه های دانش بنیان علم داده
- پوزشین های ارشد و دکتری فول فاند علم داده برای ۲۰۲۳
- فرایند استاندارد صنعتی متقاطع CRISP

## ۲- راه اندازی نرم افزاری

- زبان های برنامه نویسی علم داده
- راه اندازی پایتون در ویندوز و لینوکس
- نصب IDE های مورد نیاز
- کار با Virtual Environment
- کار با سرور لینوکس قدرتمند گوگل برای علم داده و یادگیری ماشین



Qorpi  
Engineering Workgroup

✉ qorpi

📷 qorpi.co

## مبانی محاسباتی و آنالیز

- جبر فطی
- توپولوژی و فضای ویژگی
- آمار و احتمال (میانه، میانگین، واریانس، انحراف معیار، کوواریانس)
- نظریه پیچدگی
- استقلال و هم بستگی
- علیت
- پارچوب های محاسباتی
- فرضیه سازی
- مسئله بهینه سازی

# Numpy and Scipy Framework

- انحراف معیار
- کوواریانس
- میانه
- میانگین
- همبستگی
- ضریب پیرسن
- ضریب اسپیرمن
- آرایه ها و ماتریس
- ضرب دافلی و ضرب فاربی آرایه ها
- عملیات درایه ماتریسی
- دنباله ها
- Reshape
- توزیع نرمال
- توزیع استاندارد نرمال
- اتماد و اشتراک
- مدف داده های تکراری
- واریانس

# Pandas Framework

- بارگزاری دیتاست
- کار با index
- کار با values
- Data Frame
- عملیات ها در Data Frame
- پیش پردازش اطلاعات
- پیش پردازش داده های NASA
- فرایند های آماری و آنالیز دیتاست

# Matplotlib and Seaborn Framework

- Missing values
- Normalizing
- Box plot
- Swarm plot
- Strip plot
- Pair plot
- Heatmap
- Join plot
- نمودار فطی
- نمودار نقطه ای
- نمودار میله ای
- رسم توابع
- دستورات در نمودار
- مصور سازی آنالیز داده های ناسا
- بررسی پایستگی احتمال و نرمالایز کردن
- هیستو گرام
- Pie plot
- پیش پردازش اطلاعات
- پیش پردازش اولیه داده ها

## آنالیز سری زمانی

- کانسپت های سیستم و مفهوم زمان
- مفهوم سری زمانی
- نظریه اطلاعات
- تحلیل پیوسته کرلشن
- تحلیل گسسته کرلشن
- کراس کرلشن
- معاسبه سوپرازینگ سیستم
- معاسبه چگالی بی نظمی سیستم
- تکامل سیستم
- پروژه ۱ : تحلیل زمانی بازار سهام
- پروژه ۲ : تحلیل زمانی داده های پزشکی
- پروژه ۳ : تحلیل سوپرازینگ پروژه هادرون CERN



Overfitting •

Underfitting •

روش های یاد گیری ماشین •

یاد گیری نظارت شده •

یاد گیری غیر نظارت شده •

یاد گیری نیمه نظارت شده •

یاد گیری تقویتی •

یاد گیری ماشین کوانتومی •

رگرسیون •

کلیسفیکشن •

رگرسیون فطی •

بررسی میدان مغناطیسی مول سیم حامل جریان •

گرادیان کاهش •

گرادیان کاهش برای رگرسیون فطی •

## یاد گیری ماشین

تعاریف یاد گیری ماشین •

فضای فرضیه •

تابع هزینه •

Mean Squared Error •

Mean Absolut Error •

Mean Bias Error •

Mean Squared Logarithmic Error •

Binary Cross Entropy Error •

Categorical Cross Entropy Error •

Hinge Loss •

داده های یاد گیری و داده های تست •

- Confusion Matrix •
- Classification Report •
- F1 Score •
- پروژه ۷: سیستم هوشمند تشخیص سیارک ها و شهاب سنگ های فطرناک برای زمین (NASA) •
- الگوریتم درخت تصمیم •
- الگوریتم KNN •
- بهینه سازی های پارامتر k در KNN •
- بررسی ریاضیاتی overfitting •
- Transparent Model •
- Interpretable Model and data •
- Regulation •
- Multi Object Optimization •
- Global optimum •
- Local optimum •
- تجربه یادگیری استاتیکی و دینامیکی •
- طراحی تجربه یادگیری دینامیکی •
- پروژه ۴: سیستم پیش بینی کننده بازار ارز دیجیتال •
- تست پروژه سیستم تحلیل گر و پیش بینی کننده بازار ارز دیجیتال بصورت آنلاین •
- پروژه ۵: مدل سازی دمای ستاره ها (NASA) •
- پیاده سازی و تحلیل الگوریتم رگرسیون فطی •
- Polynomial Regression •
- MLR (Multiple Linear Regression) •
- Logistic Regression •
- پیاده سازی و تحلیل الگوریتم logistic Regression •
- پروژه ۶: کلسیفیکشن انواع ستاره ها (NASA) •

- Support Vector Machine
- بردار های پشتیبان
- مرز تصمیم گیری
- Margin
- بدست آوردن مدل های فطی از مدل فطی
- افزایش بعد
- کرنل
- پروژه ۹ : استفاده از ۱۰ الگوریتم یاد گیری نظارت شده بصورت همزمان
- پروژه ۱۰ : تشخیص ضربان های نامنظم قلبی با داده ECG
- Clustering
- K-Means
- پروژه ۱۱ : فوشه بندی داده های IRIS
- فلو ت سازی مدل
- تابع جریمه یا پنالتی
- LASSO
- رگرسیون لابیستیک چند دسته ای
- شبکه های عصبی
- توابع غیر فطی فعالیت
- نورون عصبی
- Score
- Hidden Layer
- Activation Function
- SoftMax , Relu , Tanh
- پروژه ۸ : تشخیص اعداد دست نویس

- پروژه ۱۲ : تر کیب فوشه بندی با روش های نظارت شده برای کلیسیفیکشن
- کاهش ابعاد
- PCA
- بهینه سازی هایپر پارامتر تعداد components در PCA
- تر کیب PCA با الگوریتم ها نظارت شده
- پروژه ۱۳ : داده های ماهواره ای سازمان فضایی اروپا
- DBSCAN
- پروژه ۱۴ : فوشه بندی سیگنال های ECG
- بهینه سازی هایپر پارامتر های DBSCAN به کمک یاد گیری ماشین
- پاپلاین کردن الگوریتم های نظارت شده و غیرنظارت شده
- پروژ ۱۵ : پردازش تصویر و دید کشن به کمک الگوریتم های یاد گیری ماشین کلاسیک
- یاد گیری ماشین کوانتومی
- مقدمه محاسبات کوانتومی
- Superposition
- Entanglement
- Uncertainly
- Decoherence
- انواع کامپیوتر ها
- کامپیوتر های دیجیتال
- کامپیوتر های آنالوگ
- کامپیوتر های کوانتومی
- محدودیت های کامپیوتر های کوانتومی
- Quantum Gate
- Qubit