



پردیس دانشکده های فنی
دانشکده مهندسی برق و کامپیوتر

تحلیل کلان داده		نام درس
Big Data Analysis		نام درس به انگلیسی
۳ واحد	مهندسی کامپیوتر	نوع درس
کارشناسی ارشد و دکتری		مقطع
		همنیازها
		پیش نیازها
حداقل یکی از دروسهای زیر: پایگاه داده، داده کاوی، یادگیری ماشین، تشخیص الگو،		مطالب پیش نیاز
		کتاب (های) مرجع
<p>افزایش سریع خدمات آی تی باعث شده است که حجم زیادی از داده‌های بدون ساختار تولید شود. این شامل لاگ سرویس ها، سنسورها، تصاویر، صدا و غیره می‌شود. امروزه سخت‌افزار و نرم‌افزار تحلیل داده در حجم زیاد در دسترس است. اما برای تحلیل مناسب لازم است که دانشجویان با ابزار ها و تکنیک های کلان داده آشنا شوند.</p> <p>این درست مقدمه‌ای است بر مفاهیم تحلیل کلان داده شامل: data scalability, data science, cloud computing and super computing دانشجویان همچنین با تعدادی از کاربرد های کلان داده آشنا می‌شوند شامل دیتا سنتر ها، خدمات ریل تایم و دیپ لرنینگ.</p> <p>در این درس دانشجویان با تکنولوژی هایی آشنا می‌شوند که بتوانند کلان داده را در زمان مناسب دریافت و تحلیل کنند. همچنین دانشجویان با ابزار های تحلیل داده از جمله python، R، SAS آشنا می‌شوند. در این درس ابزارهای کلان داده از قبیل Hadoop، Spark و Cassandra آشنا می‌شوند. دانشجویان همچنین با مفاهیم سوپر کامپیوتر ها و مسائل عملی تحلیل داده آشنا می‌شوند.</p>		اهداف درس
Participant is expected to have a broad knowledge of tools to process big data. Participants are expected to become familiar with		نتایج درس
<ul style="list-style-type: none"> MATLAB, Python, R - Multi-threading and parallel processing - GPU programming - MapReduce and Hadoop - Elastic Search, Logstash and Kibana - Storm - 		

<p style="text-align: center;"> Message Passing Interface - Super-Computers - Parallel-algorithms - Machine Learning and Deep Learning - Real-time Internet Services - </p> <p>دانشجویانی که این درس را با موفقیت پشت سر بگذارند قابلیت‌های زیر را خواهند داشت:</p> <ol style="list-style-type: none"> ۱- ابزارهای کلان داده از قبیل Hadoop, Spark, Cassandra ۲- استک های تحلیل کلان داده از قبیل Elastic Search, Logstash, Kibana ۳- مفاهیم بنیادی کلان داده از قبیل BigTable و Dynamo و CAP Theorem ۴- سوپر کامپیوتر ها و Message Passing Interface ۵- برنامه نویسی برای GPU ۶- یادگیری ماشین و دیپ لرنینگ ۷- خدمات برخت آنلاین و نیاز آن‌ها به کلان داده و تحلیل داده 	
<ol style="list-style-type: none"> ۱- کلان داده چیست و چه وقت لازم است. ۲- چالش های کلان داده ۳- تحلیل داده با R و Python و Matlab و SAS ۴- برنامه سازی موازی و OpenMP و SIMD ۵- ارسال پیام با Message Passing Interface ۶- برنامه سازی برای GPU ۷- Hadoop و MapReduce ۸- Cassandra و Spark ۹- مقاله های BigTable و Dynamo ۱۰- دیپتا سنتر ها ۱۱- سرچ روی Cloud ۱۲- یادگیری ماشین در ابعاد بزرگ ۱۳- خدمات ریل تایم ۱۴- ترند های آینده 	<p style="text-align: center;">فهرست مباحث</p>
<p>تمام نرم افزارهای مورد نیاز Open-source و قابل دسترس هستند.</p>	<p>نرم افزار های مورد نیاز</p>
<p>دانشجویان باید یک سیستم تحلیل کلان داده بنویسند که یکی از خصوصیات زیر را داشته باشد:</p> <ol style="list-style-type: none"> ۱- به صورت ریل تایم کار کند ۲- کار یادگیری ماشین انجام بدهد ۳- به صورت distributed اجرا شود ۴- یک سرویس آنلاین ارائه دهد. 	<p>تکالیف پیشنهادی</p>
<p>یک پروژه شبیه سازی یا کار با ابزار مرتبط در راستای پیاده سازی مفاهیم مرتبط با توزیع شدگی یا کنترل در سیستم های سایبر-فیزیک با مقالات مرتبط همراه با تحلیل نتایج و ارائه موضوع.</p>	<p>پروژه های پیشنهادی</p>

تمرین	20%	نمره دهی پیشنهادی
میانترم و کوییز	30%	
امتحان پایانی	20%	
پروژه پایانی	30%	
		سایر مراجع
محمد امین صادقی		تنظیم کننده
۱۶ آذر ۹۵		تاریخ تنظیم