



بیستمین کنفرانس ملی سالانه انجمن کامپیوتر ایران  
سمپوزیوم بین المللی هوش مصنوعی و پردازش سیگنال انجمن کامپیوتر ایران

کارگاه‌های آموزشی CSICC 2015 و AISP 2015

کد کارگاه	WR1501
عنوان کارگاه	کاوش داده های عظیم
چکیده	این کارگاه الگوریتم هایی برای استخراج مدل ها و اطلاعات دیگر از مقادیر بسیار زیاد داده، معرفی می کند. تاکید این کارگاه بر روش هایی است که کارا و موثرند و قابل تعمیم برای داده های عظیم هستند. در آخر بر اساس تجربیات گروه داده کاوی دانشگاه شیراز در زمینه کاوش مجموعه داده های عظیم، ما یک رویکرد گام به گام برای حل کاربردی مسائل تجزیه و تحلیل داده های عظیم ارائه می دهیم.
زبان ارائه کارگاه	فارسی
زمان برگزاری	یک شنبه ۱۰ اسفند ۱۳۹۳ - ساعت ۸ الی ۱۲ - (۴ ساعت)
مخاطبان	دانشجویان رشته مهندسی و علوم کامپیوتر، تحلیلگران داده، کارمندان فعال در حوزه های فناوری اطلاعات
ارائه دهندگان	دانشگاه شیراز دکتر اشکان سامی - مهندس ندا ازوجی
پیش نیازهای کارگاه	آشنایی مقدماتی با مباحث داده کاوی، یادگیری ماشین
هزینه	۵۰۰۰۰۰ ریال
توضیحات	گواهی نامه حضور در کارگاه برای شرکت کنندگان صادر می شود.



بیستمین کنفرانس ملی سالانه انجمن کامپیوتر ایران  
سمپوزیوم بین المللی هوش مصنوعی و پردازش سیگنال انجمن کامپیوتر ایران

کارگاه‌های آموزشی CSICC 2015 و AISP 2015

سرفصل کارگاه:

(۱) چالش‌ها و فرصت‌ها در کاوش داده‌های عظیم.

(۲) مقدمه‌ای بر سیستم‌های توزیع فایل (DFS).

(۳) برخی از الگوریتم‌های مناسب برای کاوش داده‌های عظیم، مانند:

۱. **کاوش جریان داده‌ها:** "جریان‌ها" داده‌های ورودی به یک سیستم هستند که با یک نرخ بسیار بالا وارد سیستم می‌شوند، به عنوان نمونه‌ی این نوع جریان‌های داده‌ای می‌توان به داده‌های گرفته شده از یک ماهواره، یا دنباله‌ی کلیک‌ها برای یک سایت محبوب، اشاره کرد.
۲. **مجموعه آیت‌های مکرر:** نام دیگر آن "قواعد تلازمی" است. هدف ما یافتن آیت‌های است که به طور معمول باهم اتفاق می‌افتند. یک مثال کلاسیک آن تحلیل سبد خرید مشتری‌های یک فروشگاه است، با هدف یافتن اقلامی که به طور معمول باهم به فروش می‌رسند (مانند نوشیدنی و پوشک بچه).
۳. **خوشه بندی:** در این نوع الگوریتم‌ها هدف گروه بندی تعداد زیادی نقاط در تعداد محدودی گروه است، به طوری که نقاط موجود در هر خوشه دارای بیشترین شباهت به یکدیگر باشند و همچنین کمترین شباهت را به نقاط دیگر خوشه‌ها داشته باشند.
۴. دیگر الگوریتم‌های مرتبط.

(۴) ارائه یک مطالعه موردی از کاوش مجموعه داده‌های عظیم.