** به نام خدا سال تحصیلی:01-94**

|  |  |  |  |
| --- | --- | --- | --- |
|  **شماره تلفن:** | **شماره اتاق:**  | **نام مدرس: دکتر مهدی غضنفری**  | **دانشکده : صنایع** |
| **نام درس : انبارهای داده و داده کاوی** |
| **Email: mehdi@iust.ac.ir** | **پیشنیاز:-----** | **تعدادواحد: 3** |
| **جایگاه درس در برنامه درسی دوره :**یکی از منابع مهم کسب دانش در همه فعالیتهای بشری، انبوه داده هایی است که بتدریج در بخش های مختلف جمع آوری می شود. داده کاوی در اصل حلقه رابط میان مسایل و مشکلات از سویی و روشهای تصمیم گیری از سوی دیگر است. یعنی داده کاوی، دانش مورد نیاز برای تصمیمات درست را فراهم می کند، دانش و حکمتی که گمشده انسان است. |
| **هدف کلی** :آگاهی از انواع شیوه های مختلف کشف دانش از طریق کاویدن داده های موجود و استخراج دانش نهفته در آنها برای تصمیم گیری و بهینه سازی. دانش بدست آمده ممکن است به شکل قواعد منطقی و یا درخت های دسته بندی و یا صور دیگر بیان گشته و یا نمابش داده شود. بدیهی است برای انجام این مهم چگونگی پاکسازی داده ها و کاهش ابعاد حجم وسیع آنها باید مد نظر بوده و همچنین نرم افزار مناسبی به منظور داده کاوی بکار گرفته شود. |
| **اهداف عینی:** **اهداف مفهومی**: آشنای با مفاهیم داده، پاکسازی داده، کاهش بعد، دانش استخراجی و استنباطی، قواعد منطقی، قواعد تلازمی، خوشه بندی، دسته بندی، رگرسیون، استنتاج آماری، و کاربردهای آنها **اهداف مهارتی**: آشنایی با مسایل واقعی داده کاوی و کشف دانش و همچنین یک نرم افزار تخصصی و حرفه ای**اهداف نگرشی**: چگونگی توجه به داده های یک سازمان بعنوان منابع با ارزش دانشی و چگونگی تشخیص داده های سره از ناسره به منظور کشف دانش معتبر و بالاخره چگونگی انبارش داده ها برای حفظ گنج های ذیقمت دانشی. |
| **امکانات آموزشی**: لازم است یکی دو جلسه از کلاس در یک سایت کامپیوتری برای آموزش نرم افزار برگزار گردد.  |

**محتوای درس :**

|  |  |  |
| --- | --- | --- |
| محتویات موضوع | موضوع | هفته |
| بحث در مورد فرآیند داده کاوی به همراه انواع ابزارهای مورد استفاده در آن و همچنین ضرورت پرداختن به موضوع داده کاوی. | مفاهیم اولیه داده کاوی و کشف دانش | 1 |
| تبیین متدلوژی CRISP به همراه مثالهای کاربردی آن. | متدلوژی اجرای پروژه های داده کاوی | 2 |
| تشریح مهمترین مرحله فرآیند داده کاوی یعنی پیش پردازش داده و اشاره به زیرمراحل آن شامل پاکسازی داده، یکپارچه سازی، هموارسازی و نحوه برخورد با داده ای مغشوش، پرت و ناسازگار.  | پیش پردازش داده | 3 |
| تشریح انواع روشهای خوشه بندی داده شامل k-mean, k-mediods, SOM، مبتنی بر چگالی، سلسله مراتبی و .... | انواع روشهای خوشه بندی داده | 4 |
| معرفی انواع روشهای دسته بندی داده شامل روشهای بیزین، درخت تصمیم، شبکه عصبی، نزدیکترین همسایگی و روش های ارزیابی. | انواع روشهای دسته بندی داده | 5 |
| معرفی مفاهیم و روشهای قواعد تلازمی همچون درجه اطمینان و پشتیبان و الگوریتمهای AIS, SETM, APRIORY, …  | قواعد تلازمی | 6 |
| طرح مباحثی همچون نگاشت سریع، تحلیل مولفه اصلی، مقیاس گذاری چندبعدی، موجک گسسته، تبدیل فوریه و ... | انواع روشهای کاهش داده | 7 |
| معرفی OLAP و OLTP و همچنین مفاهیمی همچون مباحث بازارچه داده؛ پایگاه داده و انبار داده. | انباره داده | 8 |
| معرفی مفاهیم اولیه سری های زمانی و الگوریتمهای DTW و LCSS  | سری های زمانی | 9 |
| تبیین کاربرد مباحث بخشهای قبل و چگونگی استفاده از ابزارهای داده کاوی در مبحث CRM  | کاربرد داده کاوی در مدیریت ارتباط با مشتری | 10 |
| تبیین چگونگی استفاده از ابزارهای داده کاوی در مبحث Social Network  | کاربرد داده کاوی در شبکه های اجتماعی | 11 |
| معرفی مباحث نوین داده کاوی همچون روش های نوین خوشه بندی، روشهای تعیین تعداد خوشه، روشهای نوین دسته بندی، فرآیند کاوی، متن کاوی، رگرسیون لجستیک، چگونگی برخورد با کلاسهای نامتوازن، یادگیری ماشینی و ...  | مباحث پیشرفته داده کاوی | 12 |
| آموزش نرم افزار clementine، matlab و یکی از نرم افزارهای متن کاوی  | آموزش نرم افزار | 13 |
| **منابع درس** |
| - مهدی غضنفری، سمیه علیزاده و بابک تیمورپور، "داده کاوی و کشف دانش" ، انتشارات دانشگاه علم و صنعت، چاپ اول ۱۳۸۷ و چاپ دوم ۱۳۹۰- مهدی اسماعیلی "داده کاوی: مفاهیم و اگوریتم ها" 1393- سمیه علیزاده و سمیرا ملک محمدی "داده کاوی با نرم افزار" چاپ اول ۱۳۹۰ انتشارات دانشگاه خواجه نصیرالدین طوسی- [جمال شهرابی](http://www.adinehbook.com/gp/search/ref%3Dpd_sa_top/207-2552733-2638188?search-alias=books&author=%D8%AC%D9%85%D8%A7%D9%84+%D8%B4%D9%87%D8%B1%D8%A7%D8%A8%DB%8C&select-author=author-exact)، علی ذوالقدر شجاعی "داده کاوی پیشرفته مفاهیم و الگوریتم ها" جهاد دانشگاهی (دانشگاه صنعتی امیرکبیر) (1390)- [محمد صنیعی آباده](http://www.adinehbook.com/gp/search/ref%3Dpd_sa_top/207-2552733-2638188?search-alias=books&author=%D9%85%D8%AD%D9%85%D8%AF+%D8%B5%D9%86%DB%8C%D8%B9%DB%8C+%D8%A2%D8%A8%D8%A7%D8%AF%D9%87&select-author=author-exact) سینا محمودی، [محدثه طاهرپرور](http://www.adinehbook.com/gp/search/ref%3Dpd_sa_top/207-2552733-2638188?search-alias=books&author=%D9%85%D8%AD%D8%AF%D8%AB%D9%87+%D8%B7%D8%A7%D9%87%D8%B1%D9%BE%D8%B1%D9%88%D8%B1&select-author=author-exact) "داده کاوی کاربردی" نیاز دانش (1391)* Jiawei Han, Micheline Kamber, Data Mining: Concepts and Techniques, Morgan Kaufman, 2014.
* Daniel T. Larose, Discovering knowledge in Data: An Introduction to Data Mining, Wiley, 2005
 |

**شیوه ارزیابی :**

|  |  |  |
| --- | --- | --- |
| **درصد نمره** | **عنوان** | **ردیف** |
| 50% | پایان ترم | **1** |
| 25% | تمارین | **2** |
| 25% | ارائه و فعالیت کلاسی | **3** |

**وظایف دانشجویان :**

|  |  |
| --- | --- |
| **شرح وظایف** | **ردیف** |
| در راستای کسب مهارتهای پژوهشی، هر یک از دانشجویان مقطع کارشناسی ارشد ملزمند بعد از تسلط بر درس، در مورد یکی از مباحث کتاب و یا کاربردهای آن گزارشی را بصورت سمینار ارایه نمایند. دانشجویان مقطع دکتری نیز در این ارتباط می بایست یکی از مباحث پیشرفته داده کاوی را ارایه نمایند. بعلاوه آنها باید به عنوان خروجی کار خود یک فایل Powerpoint و Word در موعد مقرر تهیه و تحویل دهند . | **1** |
| به تمارین تهیه شده در موعد مقرر پاسخ داده و جواب آنها را ارسال نمایند. | **2** |
| در کلاس به هنگام تدریس مباحث مشارکت موثر داشته باشند تا به غنای مطالب افزوده گردد. | **3** |
| پروژه پایان ترم خود را با استفاده از نرم افزاری که آموزش آنرا دیده اند تهیه کرده و در موعد مقرر تحویل دهند. پروژه پایانی باید حتی الامکان بر روی داده های واقعی صورت پذیرد. | **4** |