



# کنگره بین المللی علوم و مهندسی

آلمان - هامبورگ

اسفند ماه 1396

## (طراحی انباره داده اطلاعات مهندسی نفت به کمک کاوش قواعد انجمنی)

سعید صادقی (کارشناسی ارشد مهندسی نرم افزار کامپیوتر - دانشگاه آزاد اسلامی - واحد آشتیان)

فناوری اطلاعات و ارتباطات شرکت ملی مناطق نفتخیز جنوب (سازمان حمایت کننده)

Said.sadeghi32@gmail.com

عباس کریمی (استادیار دانشکده فنی و مهندسی - گروه کارشناسی ارشد کامپیوتر - دانشگاه آزاد اسلامی اراک)

akarimi@iau-arak.ac.ir

سمیر بستانی (کارشناس صنعتی پروژه - مشارکت کننده)

فرشید احسانی (کارشناس صنعتی پروژه - مشارکت کننده)

### چکیده

در دنیای پر از اطلاعات و داده‌های سازمانی امروز روش‌های استخراج داده از اهمیت بسیار بالایی برخوردار می‌باشند که می‌توانند مفیدترین داده‌ها را برای کاربران از میان میلیون‌ها داده استخراج کنند که در قالب قوانین انجمنی ارائه شده‌اند. این قوانین زیر مجموعه داده کاوی هستند. در قوانین انجمنی با پردازش مدل‌های استخراج داده پیشین و همچنین وضعیت جاری بهترین داده‌ها پیش‌بینی و استخراج می‌شوند. ما در این پژوهش یک روش جدید برای استخراج داده با استفاده از قوانین انجمنی ارائه کردیم که نسبت به روش‌های پیشین از دقت بالاتری برخوردار می‌باشد. نتایج شبیه سازی و مقایسه نشان می‌دهد که روش پیشنهادی ما عملکرد بهتری دارد.

**واژگان کلیدی:** داده کاوی، استخراج داده، قوانین پر تکرار، پیش بینی داده

### مقدمه

با ورود به عصر اطلاعات و ارتباطات و آغاز استفاده از داده‌ها و اطلاعات به‌عنوان سرمایه‌های اصلی در حرکت علمی، اقتصادی، اجتماعی و فرهنگی جوامع، سازمان‌ها و شرکت‌های مختلف و توسعه مشارکت افراد در جهان اینترنت و ارتباطات شبکه‌ای در دنیا، دغدغه‌ای بروز پیدا کرد که از جنس همین داده‌هایی بود که همه روز و با سرعت وحشتناک در دنیا و در عرصه‌های مختلفی که فناوری اطلاعات ورود پیدا کرده بود، تولید می‌شود و آن اینکه چگونه این حجم بزرگ و متنوع داده‌ها و اطلاعات را با توجه به ساختارهایی که در فضای فناوری اطلاعات وجود دارد، می‌توان مدیریت، کنترل و پردازش کرد و از آن در جهت بهبود ساختارها و سودآوری بیشتر بهره جست؟

کاوش قوانین انجمنی یک زمینه تحقیقاتی در مبحث داده‌کاوی و در راستای کشف ارتباطات جالب و با اهمیت بین اقلام اطلاعاتی در پایگاه داده‌های بزرگ و انبارهای تراکنش می‌باشد که اخیراً تلاش‌های تحقیقاتی فراوانی را به خود اختصاص داده است از مهم‌ترین کاربردهای آن‌ها می‌توان به مسئله معروف تحلیل سبد خرید، خوشه‌بندی و کلاسه‌بندی اشاره کرد. قالب کلی قوانین انجمنی به صورت  $Y \rightarrow X$  می‌باشد که نشان‌دهنده نوعی پیش آمد هم‌زمان بین اقلام  $X$  و  $Y$  است (  $X$  و  $Y$  دو مجموعه قلم دلخواه از انباره داده‌ها هستند )  $X$  و  $Y$  به ترتیب مقدم و تالی قانون نامیده می‌شوند. معیارهای مختلفی برای