



# نقش تحلیل‌گری داده‌ها (داده‌کاوی)

## در هوشمندسازی نظارت مالی؛ کاربردها و روش‌ها

---

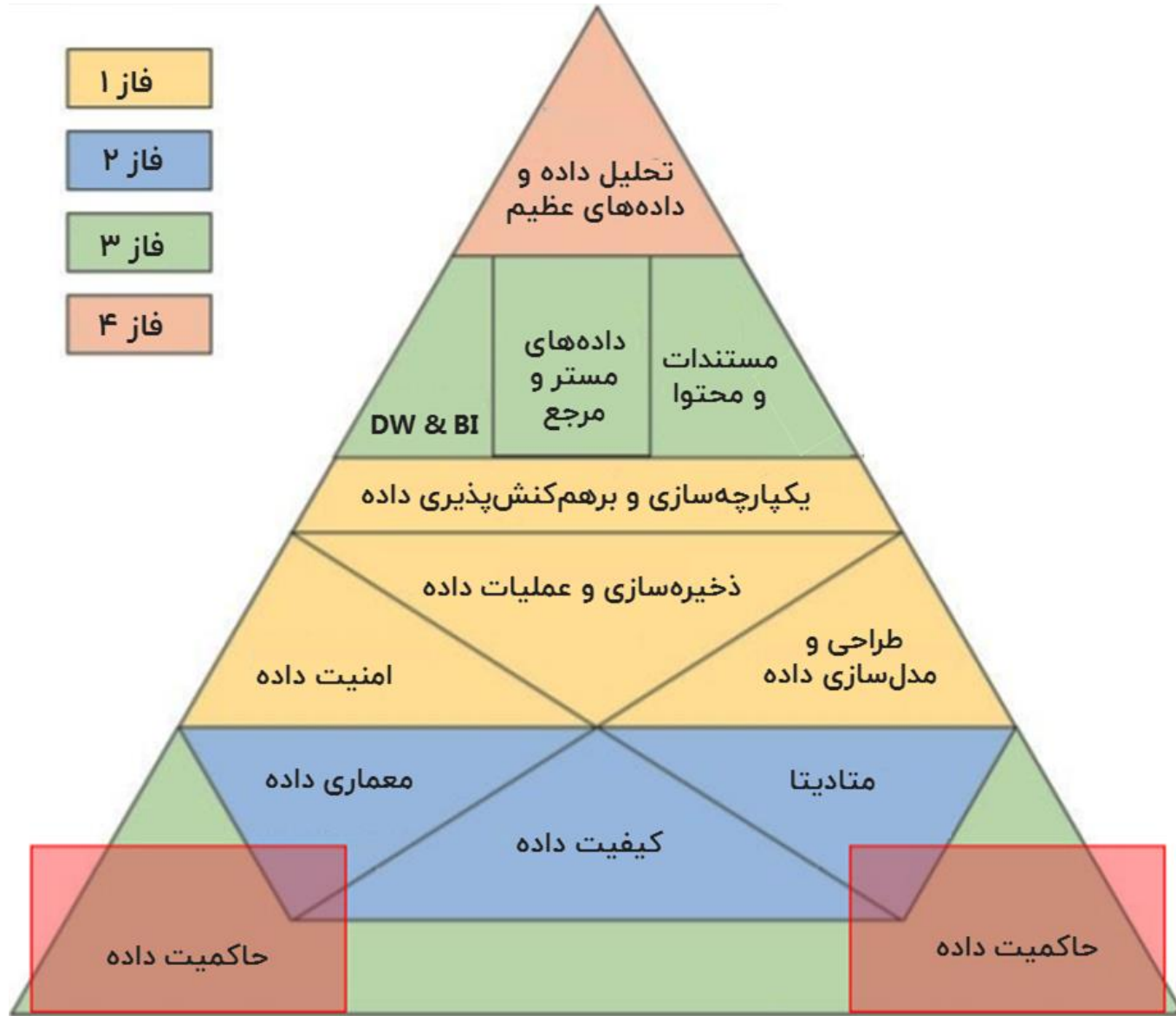
سید مهدی حسینی

کارگروه ارتقای نظارتهای هوشمند، بر خط، فناورانه و مردم پایه

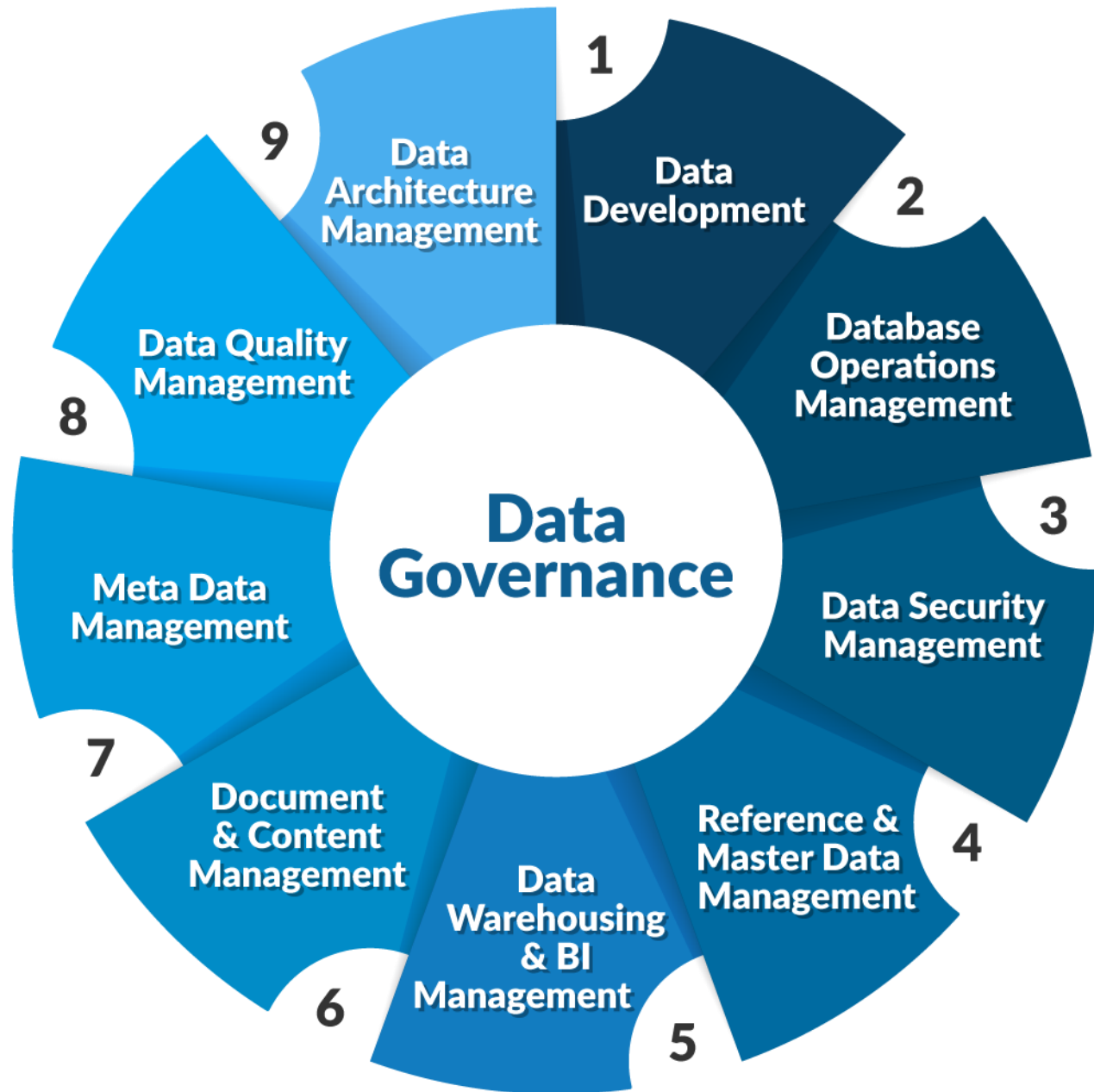
# Business Intelligence Process



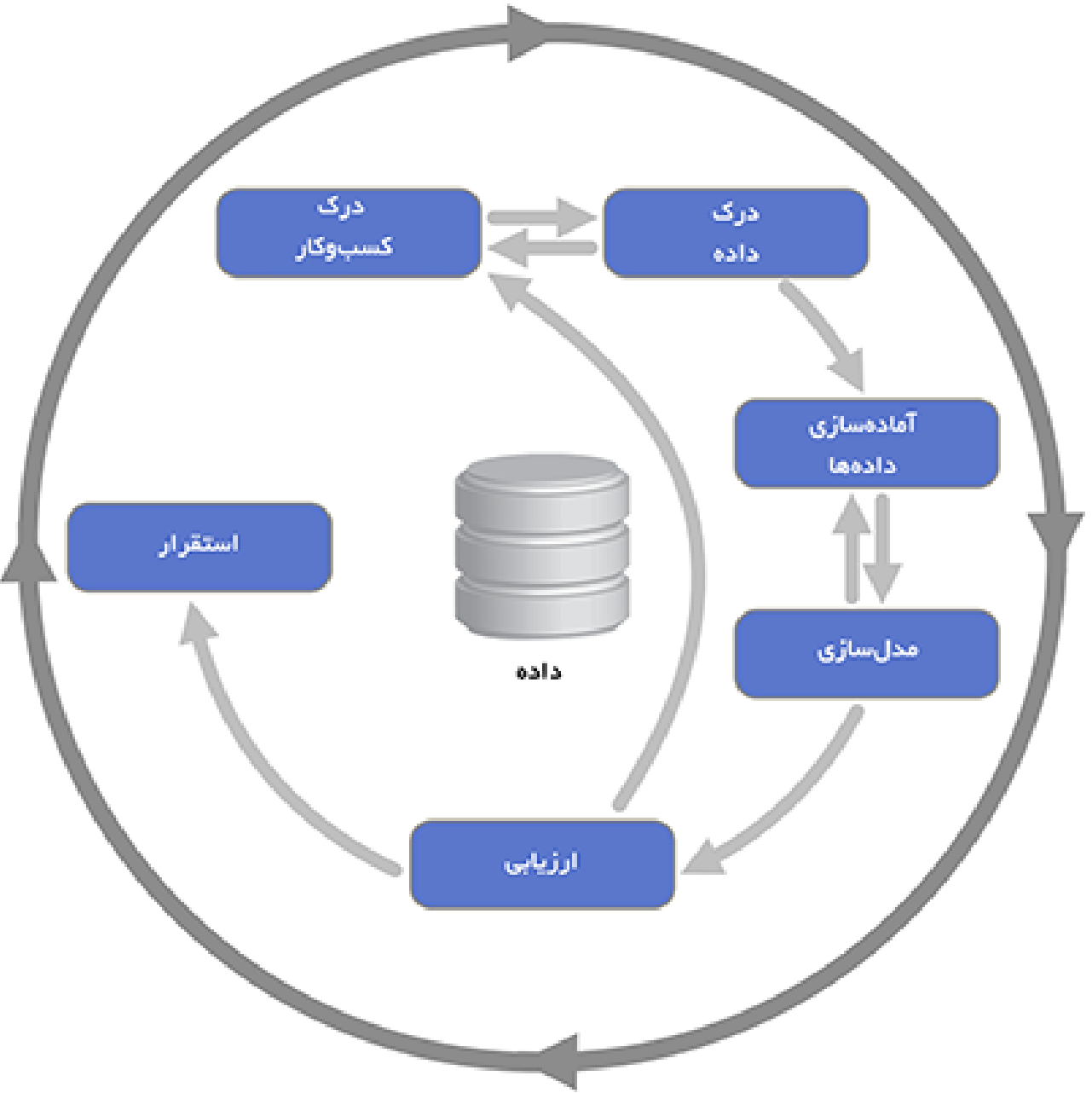
# هرم DMBOK یا مدل Aiken



# حکمرانی داده



# مراحل اصلی داده کاوی



# پاکسازی داده Data Cleaning

- اصلاح یا حذف داده‌های ناصحیح
- حذف داده‌های تکراری
- تصمیم‌گیری در مورد داده‌های ناقص
- کدگذاری صحیح داده‌ها
- ...

# تحلیل کسب و کار

توصیفی

پیش بینی

تجویزی

سوالات

چه اتفاقی افتاده است؟  
چه اتفاقی خواهد افتاد؟

چه اتفاقی خواهد افتاد؟  
چرا آن اتفاق خواهد افتاد؟

چکار باید انجام دهیم؟  
چرا آن کار را باید انجام دهیم؟

توانمندسازها

گزارشهای کسب و کار  
داشبوردها  
کارتهای امتیازی  
انبارهای داده

داده کاوی  
متن کاوی  
وب و مدیا کاوی  
پیش بینی

بهینه سازی  
شبیه سازی  
مدل سازی تصمیم گیری  
سیستم های خبره

دستاوردها

مشکلات و فرصت های کسب و  
کار به خوبی تعریف شده است

پیش بینی دقیق رویدادها  
و نتایج آینده

بهترین تصمیمات و اقدامات  
تجاری ممکن



Data Mining Tasks & Methods	Data Mining Algorithms	Learning Type
<b>Prediction</b>		
Classification	Decision Trees, Neural Networks, Support Vector Machines, kNN, Naive Bayes, GA	Supervised
Regression	Linear/Nonlinear Regression, ANN, Regression Trees, SVM, kNN, GA	Supervised
Time series	Autoregressive Methods, Averaging Methods, Exponential Smoothing, ARIMA	Supervised
<b>Association</b>		
Market-basket	Apriori, OneR, ZeroR, Eclat, GA	Unsupervised
Link analysis	Expectation Maximization, Apriori Algorithm, Graph-Based Matching	Unsupervised
Sequence analysis	Apriori Algorithm, FP-Growth, Graph-Based Matching	Unsupervised
<b>Segmentation</b>		
Clustering	k-means, Expectation Maximization (EM)	Unsupervised
Outlier analysis	k-means, Expectation Maximization (EM)	Unsupervised

**FIGURE 4.2** A Simple Taxonomy for Data Mining Tasks, Methods, and Algorithms.

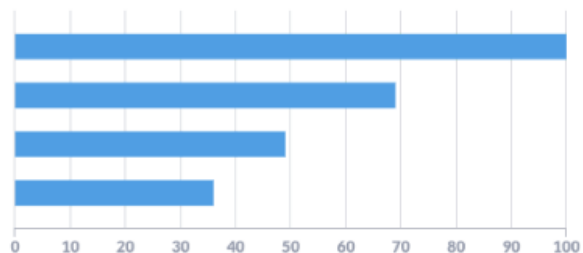




Search...

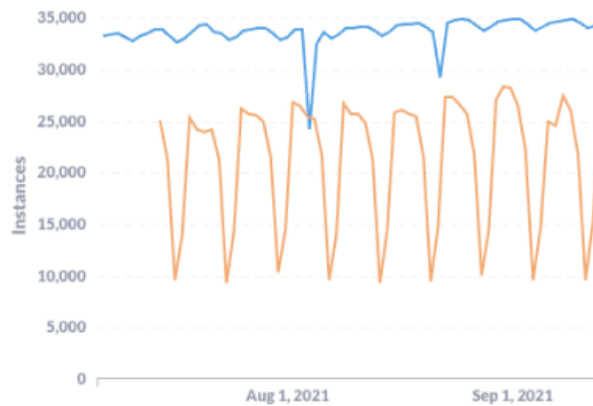
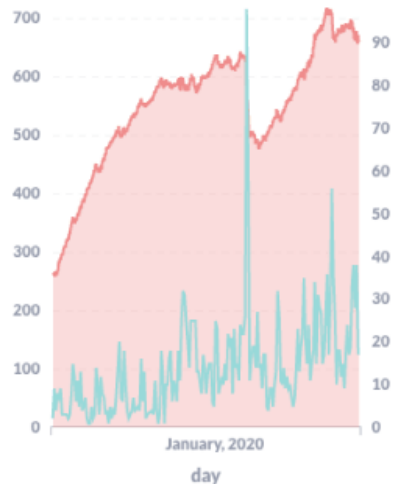
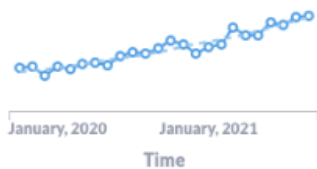
+ Ask a question

### AWS monitoring



# 35,361.5

↑ 1.6% + was 34,808.5 last day



# گزارشهای توصیفی

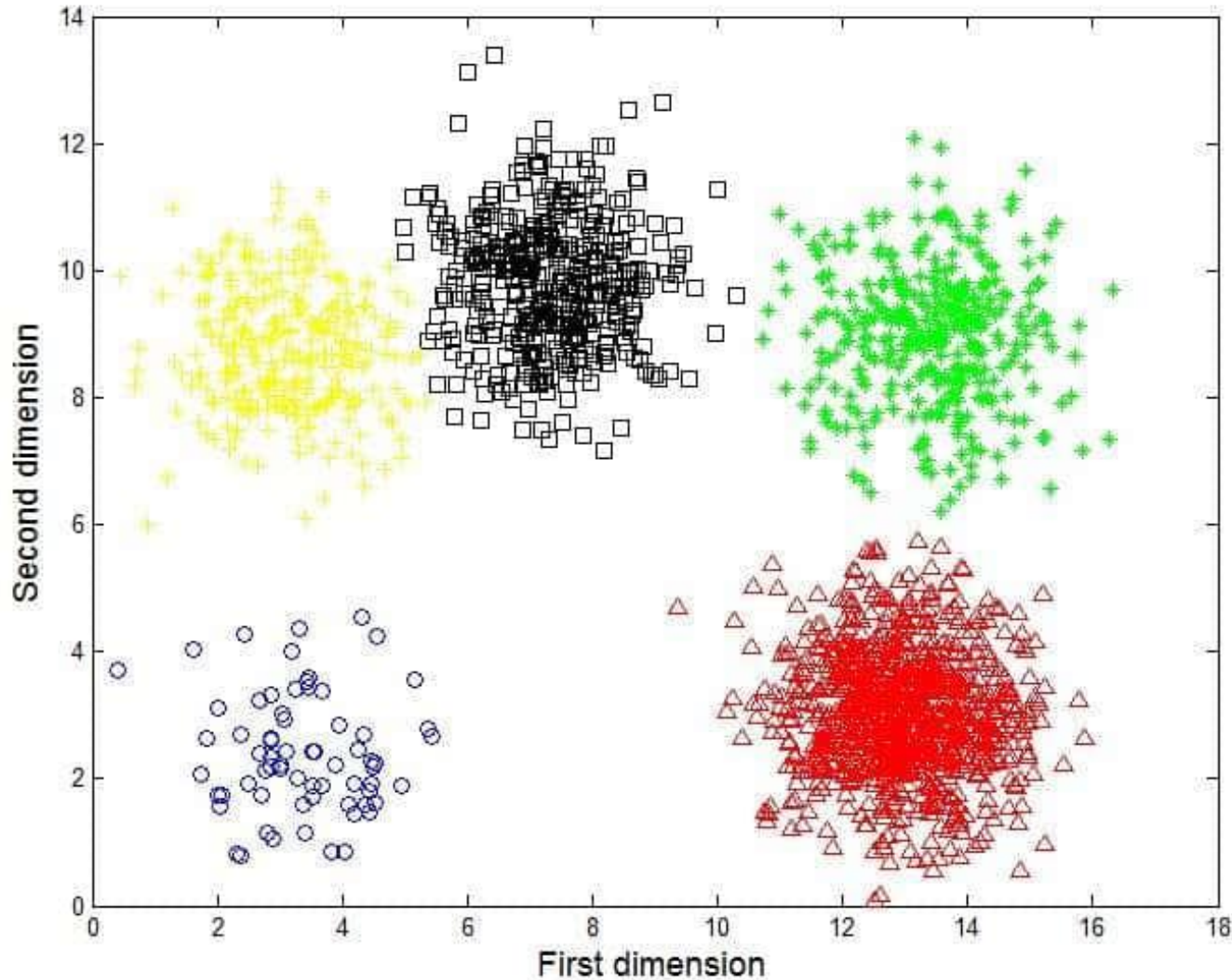
# گزارش‌های توصیفی

- «موضوعات تخلف به تفکیک قوانین و مقرره‌های قانونی»: از حیث شناسایی احکام قانونی که به صورت مکرر نقض گردیده و یا در اجرای آن ترک فعل حادث شده، مفید می‌باشد.
- «تخلفات به تفکیک دستگاه اجرایی»: از حیث شناسایی احکام قانونی که در مجموعه یک وزارتخانه، شرکت دولتی، مؤسسات انتفاعی وابسته به دولت و یا نهاد و مؤسسات عمومی غیردولتی به صورت مکرر نقض گردیده و یا در اجرای آن ترک فعل حادث شده، کاربردی است.
- «اقدامات پیشگیرانه (فواید مالی)، پیشنهاد اصلاح قوانین و مقررات، پیشنهاد اصلاح روش‌ها و دستورالعمل‌ها، اظهارنظر در ارتباط با مؤلفه‌های عملکرد، اظهارنظر در ارتباط با نظام کنترل داخلی به تفکیک دستگاه اجرایی»: از حیث شناسایی نکات حائز اهمیت در تدوین برنامه و دامنه رسیدگی مفید و حائز اهمیت است.
- «تخلفات به تفکیک مسئولین تخلف»: حیث شناسایی مدیران پر ریسک کاربردی است.
- «تخلفات به تفکیک منابع اعتباری»: از حیث شناسایی انحرافات به تفکیک اعتبارات هزینه‌ای، تملک دارایی‌های سرمایه‌ای، تملک دارایی‌های مالی و اختصاصی حائز اهمیت است.

# طبقه بندی Classification

سابقه کار	مامور یا غیر مامور	سن	دستگاه اجرایی	حقوق بالاتر از سقف دریافت کرده است؟
۱۰ تا ۱۵ سال	مامور	۳۰ تا ۴۰ سال	دستگاه X	بلی
۱۵ تا ۲۰ سال	کارمند دستگاه	۴۰ تا ۵۰ سال	دستگاه Y	خیر
۵ تا ۱۰ سال	کارمند دستگاه	۴۰ تا ۵۰ سال	دستگاه X	بلی
۱۰ تا ۱۵ سال	مامور	۴۰ تا ۵۰ سال	دستگاه Z	بلی
۱۵ تا ۲۰ سال	کارمند دستگاه	۲۰ تا ۳۰ سال	دستگاه N	خیر
۱۰ تا ۱۵ سال	کارمند دستگاه	۳۰ تا ۴۰ سال	دستگاه Y	بلی

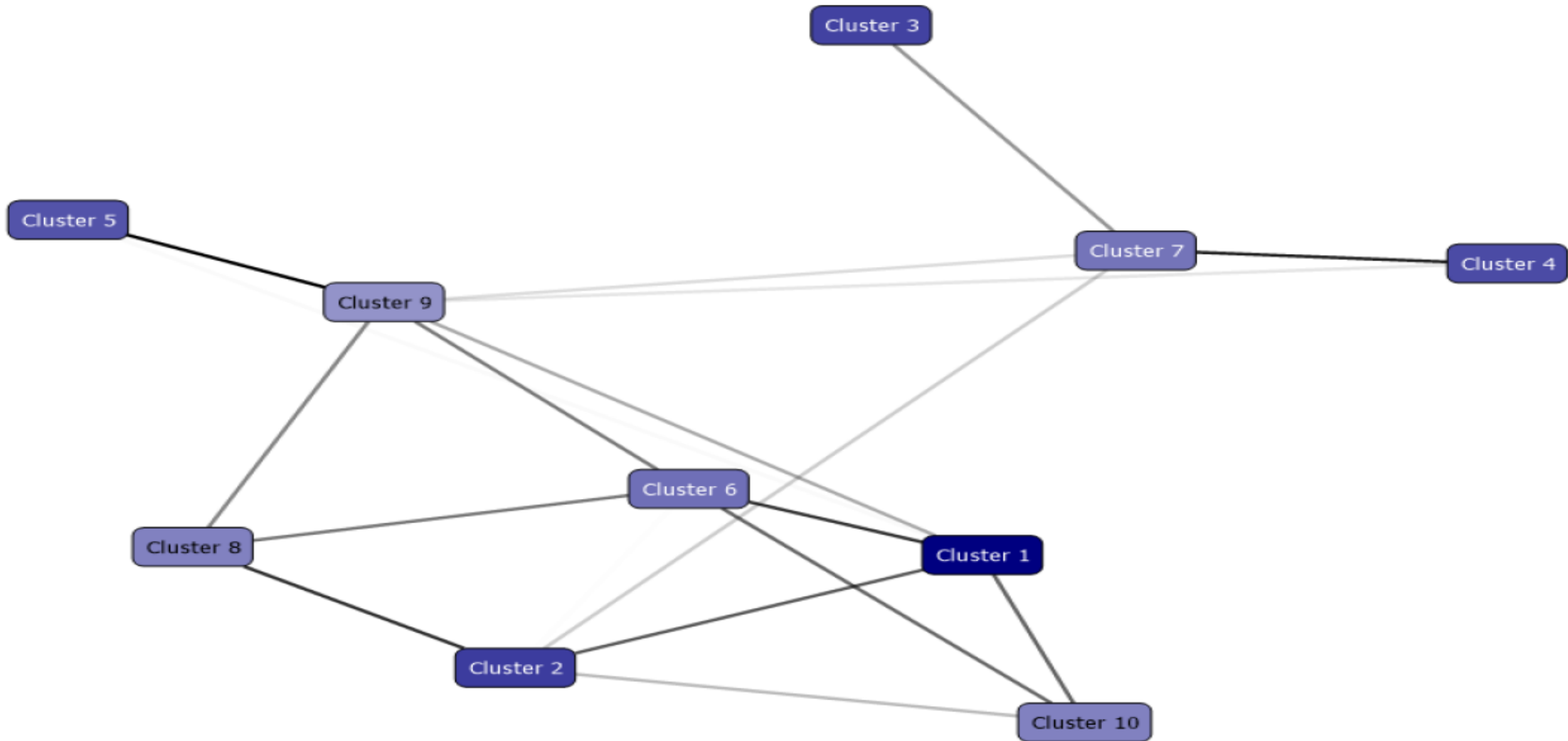
# خوشه بندی Clustering



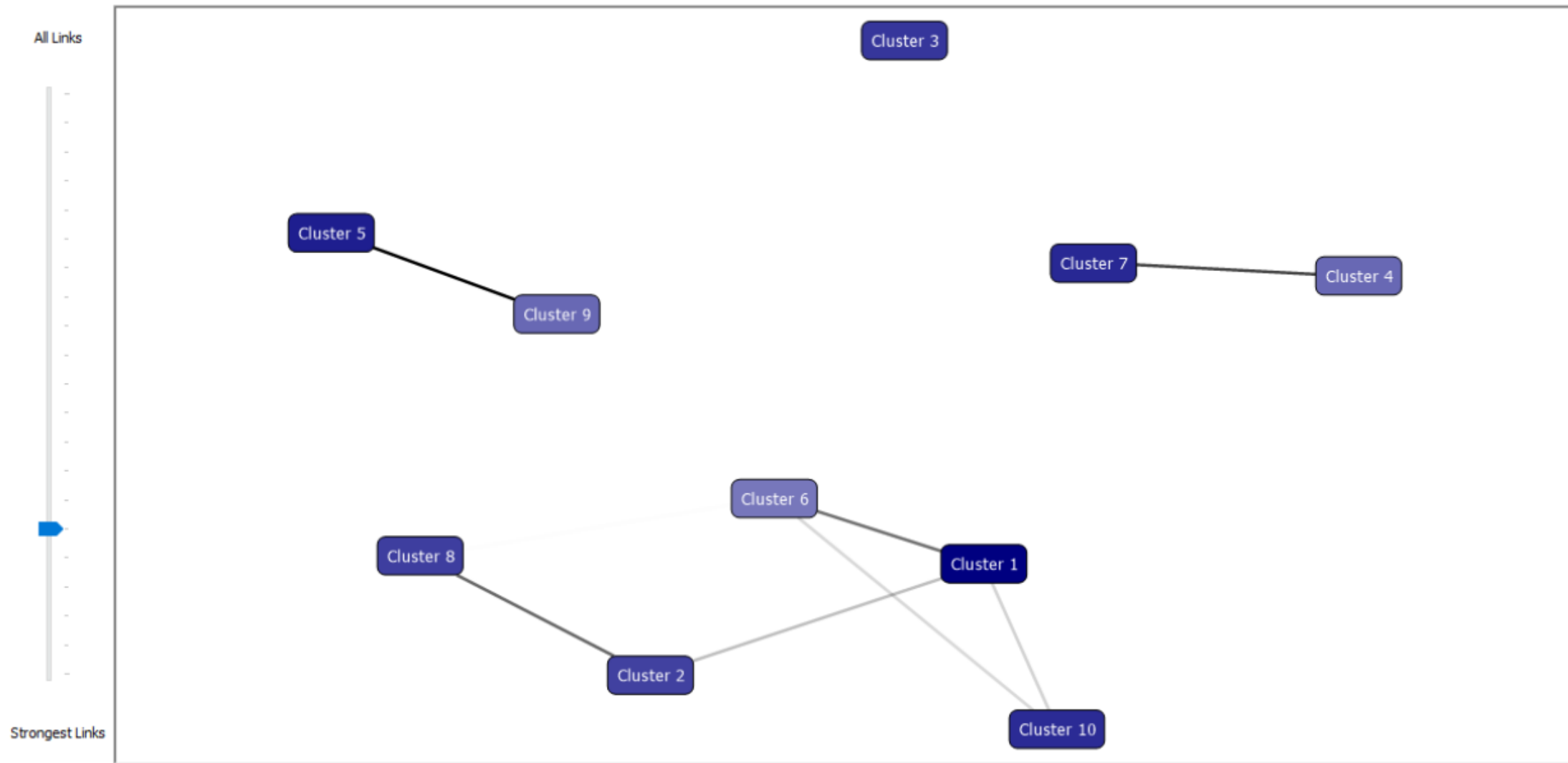
۱- از میان داده‌های دریافت‌شده، گروه‌های مختلفی از نمونه‌هایی را که شبیه به هم هستند، پیدا می‌کند. به عنوان مثال ۱۰۰ هزار داده از یک جنس داریم که آنها را به گروه‌های مشخص شده تقسیم می‌کند.

۲- تعیین ارتباط میان خوشه‌ها

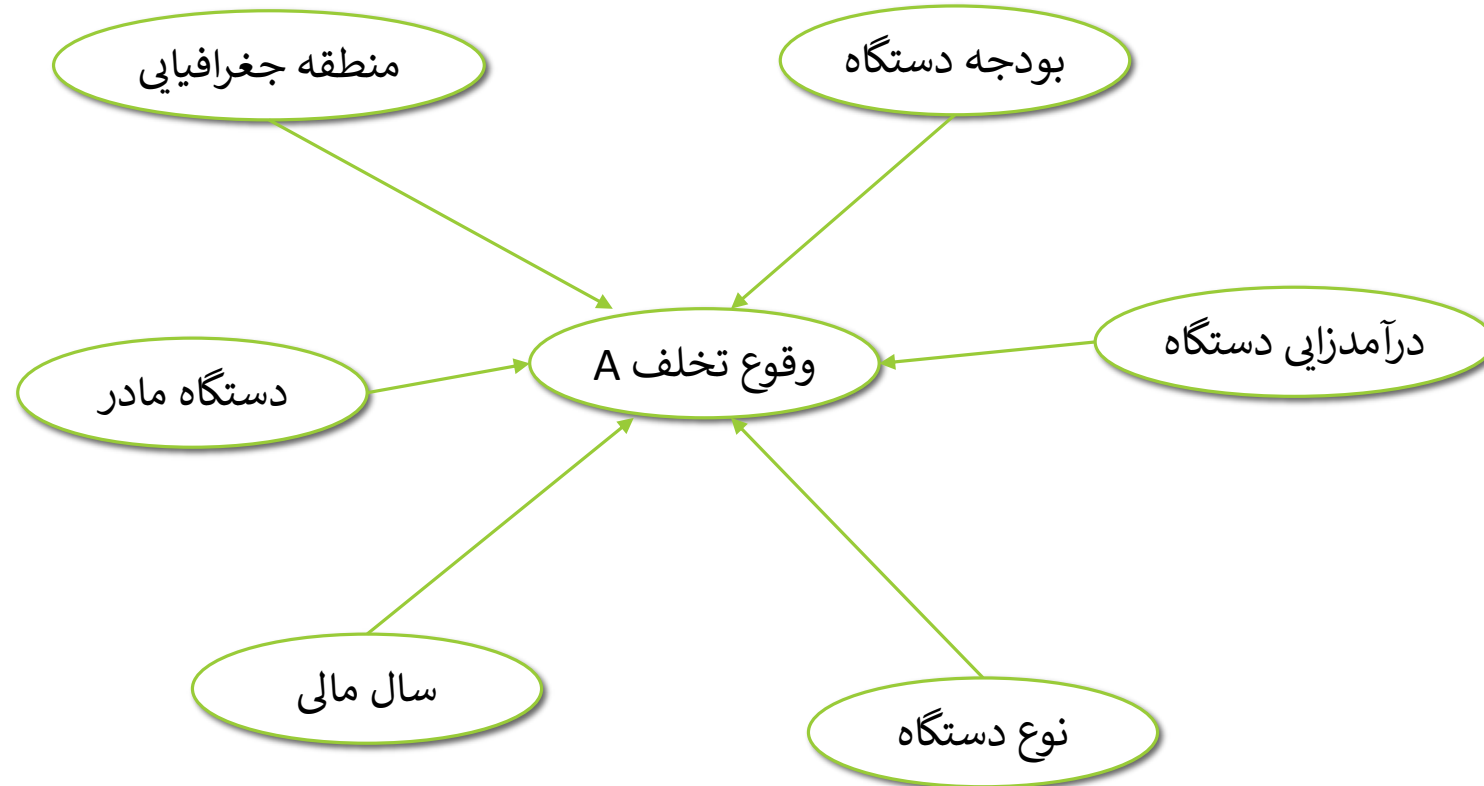
# خوشه بندی Clustering



# خوشه بندی Clustering

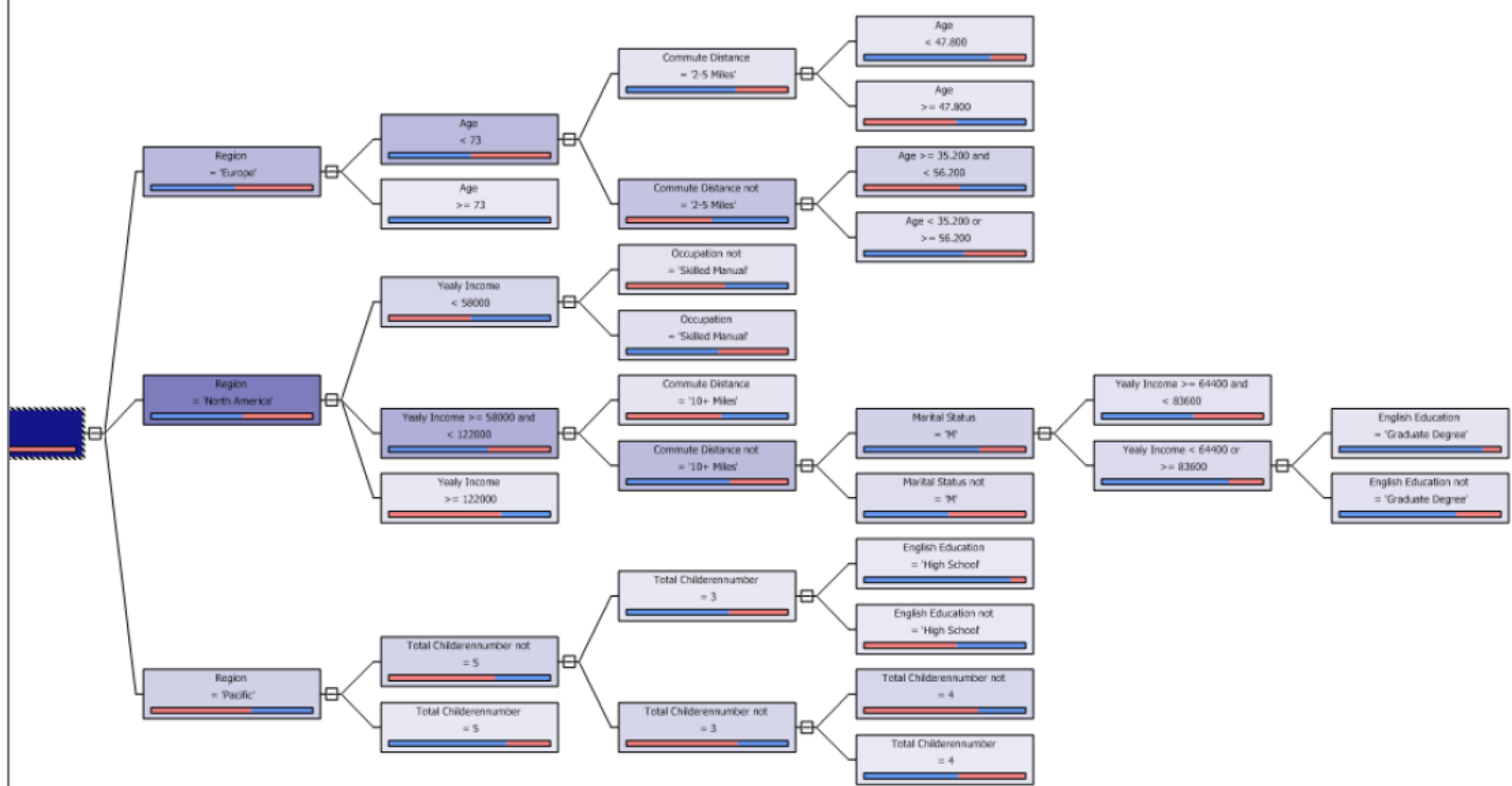


# درخت تصمیم

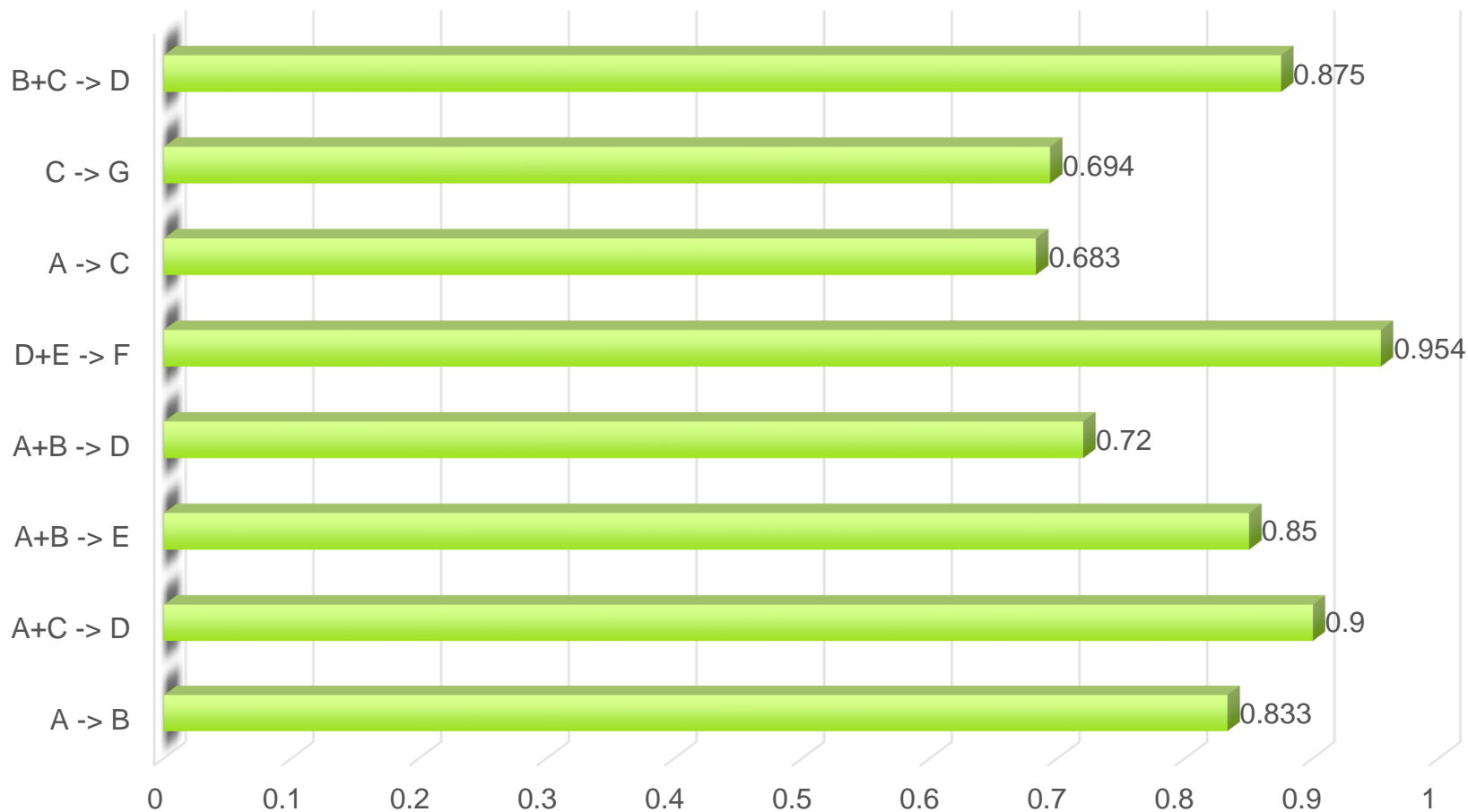




# درخت تصمیم



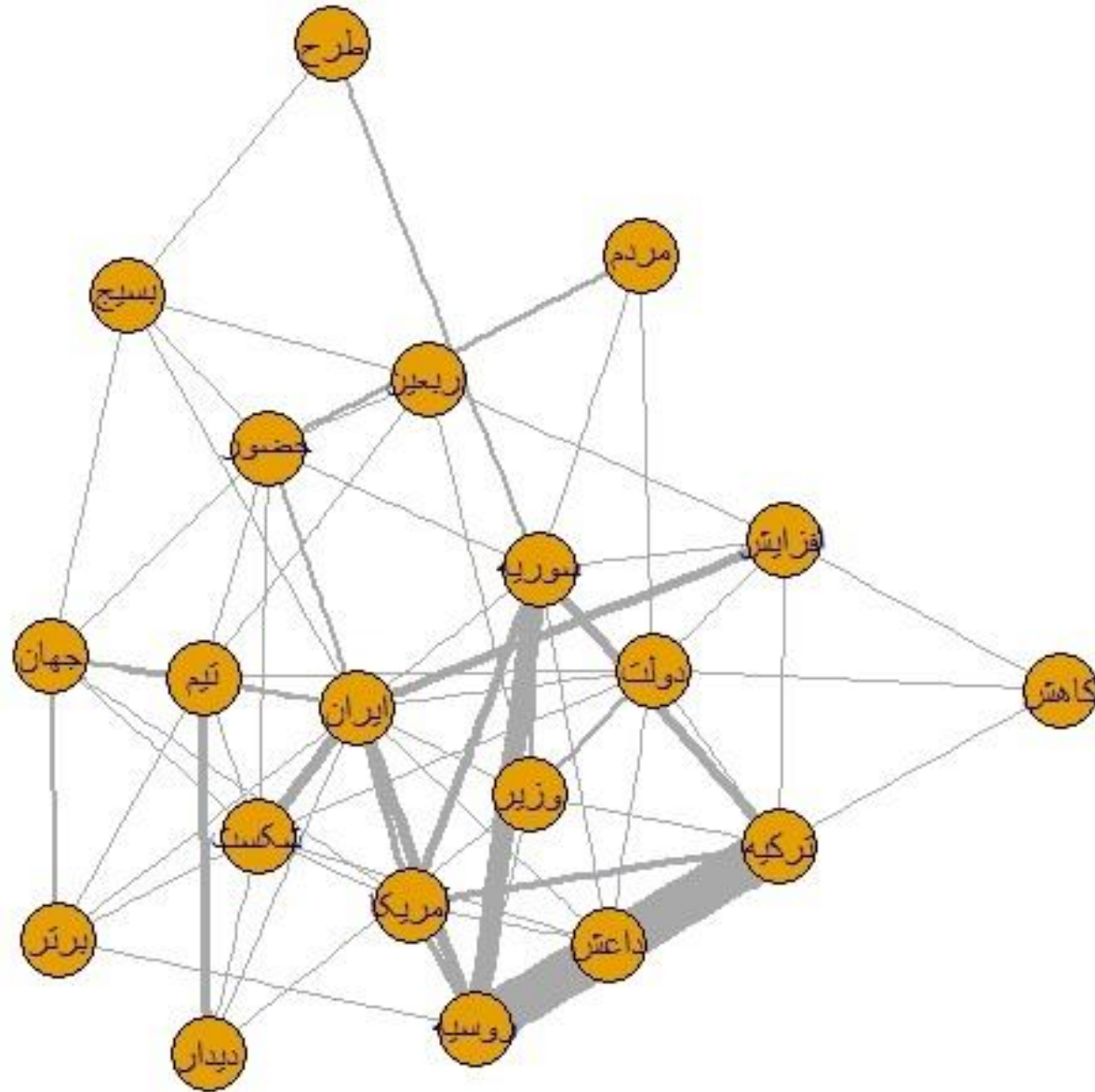
# قواعد انجمنی



# متن کاوی

افزایش <sup>نظام</sup> روز <sup>یادروزی</sup> خاتوانه <sup>عملیات</sup> تیم <sup>سیاسی</sup> خارجه <sup>لرستان</sup> <sup>بسیجیان</sup> اقتصادی <sup>تلاش</sup> عراق <sup>فروین</sup> برگزار <sup>سوریه</sup> شکست <sup>دیدار</sup> فوتبال <sup>توسعه</sup> ارائه <sup>صهیونیستی</sup> حضور <sup>پاکستان</sup> تهران <sup>بسیج</sup> <sup>تروریسم</sup> یمن <sup>انتخابات</sup> آماده <sup>فرانسه</sup> سال <sup>وزارت</sup> وارد <sup>جامعه</sup> علمی <sup>دانشگاه</sup> توسط <sup>تهدید</sup> کشته <sup>ادامه</sup> لیگ <sup>منطقه</sup> آمریکا <sup>ایجاد</sup> امام <sup>آینده</sup> <sup>تولید</sup> <sup>نشان</sup> رئیس <sup>نقود</sup> <sup>جهانی</sup> برابر <sup>اصلی</sup> پژوهشی <sup>مجلس</sup> <sup>پایتان</sup> تحریم <sup>همه</sup> <sup>زائران</sup> <sup>میان</sup> آموزش <sup>زیر</sup> <sup>هواپی</sup> برنامه <sup>مرز</sup> داعش <sup>مبارزه</sup> کشتی <sup>دشمن</sup> <sup>حمایت</sup> مرحله <sup>پرونده</sup> <sup>عربستان</sup> فرهنگی <sup>رژیم</sup> لزوم <sup>برخورد</sup> ملی <sup>فعالیت</sup> روسی <sup>برتر</sup> مشکلات <sup>گزارش</sup> اردوستان <sup>اعلام</sup> <sup>شهرداری</sup> هزینه <sup>قتال</sup> <sup>زائر</sup> اسلامی <sup>تیملی</sup> سفر <sup>حال</sup> <sup>اربعین</sup> <sup>نجات</sup> بیشتر <sup>هفته</sup> <sup>طرح</sup> گلستان <sup>رنال</sup> <sup>ایران</sup> <sup>تشکیل</sup> <sup>بسجی</sup> <sup>واکنش</sup>

# متن کاوی



# سیستمهای توصیه گر



# آزمون قواعد

آیا بدهی مربوطه ثبت شده است؟

شرکت طرف قرارداد ثبت شده است؟

موافقتنامه مربوطه مورد تایید است؟

مبلغ پرداختی با سیستم حسابداری مطابقت دارد؟

طرف قرارداد کارمند دولت است؟

مناقصه در سامانه ستاد ثبت شده است؟

هویت طرف قرارداد احراز می شود؟

مبلغ مندرج در سند با مبلغ پرداختی تطبیق دارد؟

## قرارداد

سامانه هوشمند حسابرسی

سامانه سماد

سامانه ثبت شرکتها

سامانه بودجه

سامانه خزانه

سامانه پاکنا

سامانه ستاد

سامانه ثبت احوال

سامانه حسابداری دستگاه اجرایی

# از توجه شما سپاسگزاریم

سید مهدی حسینی

کارگروه ارتقای نظارتهای هوشمند،

بر خط، فناورانه و مردم پایه