

عنوان مقاله:

مروری بر روش های بهبود Gain Ratio در الگوریتم C4.5

محل انتشار:

دومین کنفرانس ملی رویکردهای نوین در مهندسی کامپیوتر و برق (سال: 1395)

تعداد صفحات اصل مقاله: 5

نویسندگان:

سلیمه ضیالالدینی - دانشجوی کارشناسی ارشد، مهندسی کامپیوتر (نرم افزار)، دانشگاه غیرانتفاعی بعثت کرمان

مصطفی قاضی زاده احسائی - استادیار بخش مهندسی کامپیوتر، دانشکده فنی و مهندسی، دانشگاه شهید باهنر کرمان

خلاصه مقاله:

داده کاوی به معنای استخراج اطلاعات، دانش و کشف الگوهای پنهان از یک پایگاه داده های بسیار بزرگ است که این الگوها و دانش ها معمولاً مستتر در داده می باشند. یکی از معروف ترین تکنیک های داده کاوی، درخت تصمیم است. درخت تصمیم یکی از متداول ترین روش های تقسیم بندی در زمینه داده کاوی است. در حال حاضر تعداد بسیار زیادی از الگوریتم های درخت تصمیم وجود دارد که یکی از گسترده ترین آن ها، الگوریتم C4.5 است که به دلیل کارایی و استفاده در کاربردهای داده کاوی به طور مکرر مورد استفاده قرار گرفته است. اینکه کدام خصوصیت برای قرار گرفتن در گره ریشه درخت یا زیر درخت قرار داده شود، نیاز به تعیین یک معیار مناسب دارد. در الگوریتم C4.5 از معیار Gain Ratio برای انتخاب خصوصیت در تولید درخت تصمیم استفاده می شود. لذا مطالعات پیشین، روش ها و راهکارهایی را مبنی بر بهبود بخشیدن آن جهت افزایش دقت و کاهش پیچیدگی زمان درخت تصمیم ارائه داده اند. در این مقاله، همواره سعی بر این است که به بررسی و مرور روش ها و راهکارهای ارائه شده پرداخته شود و دقت پیچیدگی زمانی به دست آمده از روش ها و راهکارهای مطالعات پیشین را، مورد بررسی قرار دهد.

کلمات کلیدی:

الگوریتم C4.5، پیچیدگی زمان، خصوصیت، داده کاوی، درخت تصمیم، دقت، Gain Ratio

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/522539>

