# محبوب‌ترین فیلم ها و سریال های وبسایت ای ام دی بی

سینا طیبی

دانشجوی علوم کامپیوتر دانشگاه شهید بهشتی

sinaapolo969@gmail.com

آریانا آقامحمدی ‌قاضیانی

دانشجوی علوم کامپیوتر دانشگاه شهید بهشتی

cogokeye@gmail.com

دیبا الهی

دانشجوی علوم کامپیوتر دانشگاه شهید بهشتی

[elahidiba82@gmail.com](mailto:elahidiba82@gmail.com)

علی بخشش

دانشجوی علوم کامپیوتر دانشگاه شهید بهشتی

bkhshshly18@gmail.com

چکیده

در این مقاله به تحلیل داده های وب‌سایت IMDb در مورد فیلم ها و سریال ها می‌پردازیم و تاثیرات هرکدام از متغییر های وابسته را بر روی متغییر های اصلی تحلیل و بررسی می‌کنیم. متغییر هایی که در جدول داده ها در وب‌سایت Kaggle موجود است، شامل دو بخش داده های مستقل و وابسته می‌شوند. داده های مستقل و اصلی ما شامل برآورد (Rate) و تعداد رأی Votes)) می‌باشد. به طبع هر چه تعداد رأی دهندگان بیشتر باشد، برآورد از این داده ها از صحت بیشتری برخوردار خواهد بود و معتبرتر است. داده هایی مانند میزان خشم مورد نمایش در فیلم ((Violence، ژانر فیلم (Genre)، مدت زمان آن (Duration)، ترسناک بودن آن (Frightening)، میزان برهنگی (Nudity)، استفاده از الفاظ رکیک (Profanity)، نشان دادن الکل و مصرف آن در فیلم (Alcohol)، نوع گواهی آن فیلم (Certificate) و علاوه بر موارد ذکر شده، برای سریال ها تعداد قسمت های آن (Episodes) است. در ادامه نوع (Type) به معنی سریال یا فیلم سینمایی بودن آن نیز می‌تواند تاثیرگذار باشد.

کلمات کلیدی : IMDb، برآورد، رأی، خشونت، ژانر، مدت‌زمان، وحشت، برهنگی.

مقدمه

بیان پرسش ها