

## الگوریتم PMI (اطلاعات متقابل نقطه‌ای)

### معیاری جهت تشخیص شباهت معنایی کلمات

#### چکیده

در فرایند استخراج و بازیابی اطلاعات، تشابه‌یابی داده‌های موجود در فضای مجازی از اهمیت خاصی برخوردار است. در پژوهش حاضر، از الگوریتم اطلاعات متقابل نقطه‌ای (PMI) جهت تشخیص شباهت معنایی جفت واژه‌ها استفاده شده است. داده‌های آماری این پژوهش از جستارهای موتور جستجوی وب بدست آمده است. جهت سنجش درصد درستی الگوریتم از نمونه سوالات چهار گزینه‌ای کلمات مترادف تافل استفاده شده و مقادیر بدست آمده از ترجمه فارسی سوالات، با نمونه مشابه انگلیسی آن مقایسه شده است. دقت عملکرد این الگوریتم در تشخیص شباهت معنایی جفت واژه‌ها در نمونه مشابه انگلیسی 74 درصد و در پژوهش حاضر 60 درصد بوده است.

**کلید واژه‌ها:** الگوریتم اطلاعات متقابل نقطه‌ای، اطلاعات مشترک، استخراج اطلاعات، احتمال شرطی،

هم‌معنایی