

۲۱۴۳۴۴۹

یافتن متون مشابه با داده کاوی

مهندس علی کرام‌الدینی
دکتر حسین ازگویی



۱۳۹۹

سرشناسه	: کرام‌الدینی، علی، ۱۳۷۱-
عنوان و نام پدیدآور	: یافتن متون مشابه با داده‌کاوی / علی کرام‌الدینی، حسین ازگومی.
مشخصات نشر	: تهران: نوشتار، ۱۳۹۹.
مشخصات ظاهری	: ۹۶ ص.
شابک	: 978-622-973540-4
وضعیت فهرست نویسی	: فیپا
یادداشت	: کتابنامه: ص. ۹۳-۹۵.
موضوع	: داده‌کاوی -- نمونه‌پژوهی
موضوع	: Data mining -- Case studies
شناسه افزوده	: ازگومی، حسین، ۱۳۴۷-
شناسه خزود	: Azgomi, Hossein
رده بندی کنگره	: Q476/9
رده بندی دیویی	: ۰۰۶۱۳
شماره کتابشناسی ملی	: ۷۳۱۶۶۹۲

یافتن متون مشابه با داده‌کاوی

مؤلفان: مهندس علی کرام‌الدینی و دکتر حسین ازگومی
طراح جلد: فریده نعمت‌اللهی
ناشر: انتشارات نوشتار
نوبت چاپ: اول، ۱۳۹۹

مجموعه: مان: ۱۰۰ نسخه
بها: ۳۰۰۰ تومان
شابک: ۹۷۸-۶۲۲-۹۷۳۵۴۰-۴
ISBN: 978-622-97354-0-4

کلیه حقوق این اثر محفوظ است.

تکثیر، انتشار و یا ذخیره‌سازی تمام یا بخشی از این اثر به هر شکل و یا هر هدف بدون مجوز از ناشر، غیرقانونی و قابل پیگرد است.

www.NeveshtarBooks.ir
NeveshtarBooks@gmail.com



فهرست مطالب

۹	فصل ۱: مقدمه و کلیات پژوهش
۱۱	۱-۱- مقدمه
۱۲	۱-۲- بیان مسئله
۱۳	۱-۳- کاربردهای مسئله یا تنگنای مشابه
۱۴	۱-۳-۱- شناسایی سرفصل‌های ادبیات
۱۴	۱-۳-۲- تشخیص صفحات آبی در دست
۱۵	۱-۳-۳- تشخیص مقالاتی از منابع یحسان
۱۶	۱-۳-۴- فیلترهای همکاری
۱۶	۱-۳-۵- خرید اینترنتی
۱۷	۱-۳-۶- رتبه‌بندی فیلم‌ها
۱۸	۱-۴- تعاریف پایه‌ای
۱۸	۱-۴-۱- تشابه جاکارد
۲۰	۱-۴-۲- درهم‌سازی
۲۳	۱-۵- نوآوری پژوهش
۲۴	۱-۶- اهمیت و ضرورت تحقیق
۲۵	۱-۷- اهداف تحقیق
۲۵	۱-۸- فرضیه‌های تحقیق
۲۶	۱-۹- سؤال‌های تحقیق

فصل ۲: مبانی نظری و پیشینه پژوهش ۲۷

۲-۱-۱- مقدمه ۲۹

۲-۲- تعاریف، اصول و مبانی نظری ۲۹

۲-۲-۱- داده‌کاوی ۳۰

۲-۲-۱-۱- فرآیند داده‌کاوی ۳۱

۲-۲-۱-۲- کاربردهای داده‌کاوی ۳۴

۲-۲-۲- متن‌کاوی ۳۵

۳-۲- معرفی بر ادبیات موضوع ۳۹

۳-۱- مرور روش‌های مبتنی بر محتوی ۴۰

۴-۲- روش‌های شاخص ۴۷

۴-۱- روش‌های ششگانه ۴۸

۴-۲-۲- روش هم‌سازگی حداقلی ۵۰

۴-۲-۳- روش هم‌سازگی ساس به مکان ۵۲

۴-۲-۴- روش تلفیقی ۵۳

فصل ۳: روش پیشنهادی ۵۵

۳-۱- مقدمه ۵۷

۳-۲- روش پیشنهادی ۵۷

۳-۲-۱- پیش‌پردازش‌ها ۵۹

۳-۲-۲- تبدیل‌ها به مجموعه ۵۹

۳-۲-۳- امتیازدهی به سندها ۶۰

۳-۲-۴- دسته‌بندی سندها ۶۱

۳-۲-۵- یافتن سندهای کاندید ۶۲

۳-۲-۶- یافتن سندهای مشابه ۶۳

۳-۳- الگوریتم روش پیشنهادی ۶۴

۳-۴- مثالی از روش پیشنهادی ۶۶

۷۱	فصل ۴: نتایج تجربی
۷۳	۴-۱- مقدمه
۷۳	۴-۲- محیط ارزیابی
۷۴	۴-۳- مجموعه داده
۷۶	۴-۴- روش‌ها
۷۶	۴-۴-۱- روش LSH
۷۷	۴-۴-۲- روش HBD
۷۷	۴-۴-۳- روش LBS
۷۷	۴-۵- ارزیابی‌های عملی
۷۷	۴-۵-۱- ارزیابی‌های امنیتی
۸۱	۴-۵-۲- ارزیابی از نظر امنیت کاذب
۸۲	۴-۵-۳- ارزیابی از نظر نفی کاذب
۸۳	۴-۶- مقایسه نتایج
۸۷	فصل ۵: نتیجه‌گیری و پیشنهادها
۸۹	۵-۱- نتیجه‌گیری
۹۱	۵-۲- پیشنهادها
۹۳	مراجع