

سازدخواست
۱۵۲۳۷۳۰



هوش مصنوعی و پردازش زبان و فرهنگ ایرانی

نویسنده:

امیرشهاب شاهمیری

ناشر:

کانون نشر علوم



کانون نشر علوم

www.nashreloom.com

پیشگام در نشر آثار برگزیده علوم کاربردی کامپیوتر و الکترونیک

سرشناسه: شاهمیری، امیرشهاب، ۱۳۵۳ -

عنوان و نام پدیدآور: هوش مصنوعی و پردازش زبان و فرهنگ ایرانی / مؤلف امیرشهاب شاهمیری. **

مشخصات نشر: تهران: کانون نشر علوم، ۱۳۹۶.

مشخصات ظاهری: ۲۰۶ ص: مصور، جدول، نمودار.

شابک: ۹۷۸-۹۶۴-۳۲۷-۱۸۰۰

وضعیت فهرست نویسی: فیا

یادداشت: ص.ع. به انگلیسی: Amir Shahab Shahmiri. Iranian culture and language processing using artificial intelligence.

یادداشت: واژه نامه.

یادداشت: کتابها: ص. ۹۰

یادداشت: نمایه.

موضوع: پردازش زبان طبیعی - مقاله ها و خطابه ها

موضوع: Addresses, essays, lectures --* Natural language processing (Computer science)

موضوع: هوش مصنوعی - مقاله ها و خطابه ها

موضوع: Addresses, essays, lectures --* artificial intelligence

موضوع: الگوریتم های ژنتیک - مقاله ها و خطابه ها

موضوع: Addresses, essays, lectures --* Genetic Algorithms

رده بندی کنگره: ۱۳۹۶ ش ۲/۴ پ ۷۶/۹ QA

رده بندی دیویی: ۰۰۶/۳۵

شماره کتابشناسی ملی: ۴۸۲۶۴۷۶

نام کتاب: هوش مصنوعی و پردازش زبان و فرهنگ ایرانی

ناشر: کانون نشر علوم

تألیف: امیرشهاب شاهمیری

صفحه آرا: فاطمه آدیگوزل پور اردبیلی

طرح جلد: بهنام شاه حسینی

چاپ اول: پاییز ۱۳۹۶

تیراژ: ۱۰۰۰

چاپ و صحافی: کانون نشر علوم

قیمت: ۱۶۰۰۰ تومان

شابک: ۹۷۸۹۶۴۳۲۷۱۸۰۰

دفتر مرکزی پخش: خیابان انقلاب، خیابان فخررازی، خیابان وحید نظری شرقی، پلاک ۶۵، واحد ۱

تلفن: ۶۶۹۶۱۵۶۸-۶۶۴۹۴۹۱۱



ناشر سخن

به نام خداوند بخشنده و مهربان

گستره دانش بشری در برابر علم بی‌پایان و نامتناهی حضرت حق (جلّ جلاله) بسیار ناچیز است؛ همچنان به کربم و ما عطا نکردیم به شما علم را جز اندکی» شاهدهی است بر ضعف معرفت انسانی نسبت به حکمت و صفاتی. با این حال چه نیکو سفارش فرموده است خاتم انبیاء، حضرت محمد مصطفی (ص) که «چه بد علم را اگر چه به سرزمین چین». این توصیه، تمامی پیروان راستین دین حنیف اسلام، به فراگیری و اکتساب علوم و فنون مختلف در همه زمان‌ها و همه مکان‌ها فرا می‌خواند.

تلاش کانون نشر علوم، همواره معطوف به تولید آثار دانشی و بایسته در زمینه فنی و مهندسی بوده است؛ اگرچه که در این راه مشکلات و دشواری‌های فراوانی حادث شده و خواهد شد. این امر تا حدّی بسیار، تابعی از ماهیت رشته‌های فنی و مهندسی و معارف مربوط به آنهاست؛ چرا که این رشته‌ها به سرعت و بی‌وقفه در حال نو به نو شدن می‌باشند. این امر رسالت کانون نشر علوم را در نشر آثار مربوطه، اعم از تألیف یا ترجمه، خطیب کرده می‌کند. در این ارتباط دغدغه اصلی کانون نشر علوم این بوده و هست که اولاً، از قافله علوم نشر مهندسی عقب نماند؛ و ثانیاً، آثاری را به زیور طبع بیاراید که ضمن داشتن وجاهت علمی، در خور تأمل مخاطبین اندیشمند و قاضل نیز باشد. امید است که خداوند متعال در این مهم مساعدت فرماید.

بایان سخن اینکه، کانون نشر علوم دست یاری و همت تمامی مؤلفین و مترجمین علاقه‌مند در زمینه علوم فنی و مهندسی را به گرمی می‌فشارد و تمامی ایشان را خاضعانه و خاشعانه به همکاری فرا می‌خواند. همچنین این مجموعه از تمامی مخاطبین استدعا دارد که با نظرات صائب و راهگشای خود، در بهبود شکلی و محتوایی آثار منتشر شده مساعدت فرمایند؛ و صد البته که تمامی کاستی‌ها از این رهگذر تنها و تنها متوجه این مجموعه بوده و هست.

سید محمدحسین منوچری

کانون نشر علوم



فهرست مطالب

۹	فصل ۱: پیشگفتار
۱۲	بشت برده مقالات
۲۱	فصل ۲: شناسایی اشعار شاهنامه فردوسی به کمک شبکه عصبی مصنوعی
۲۲	۱- مقدمه
۲۳	۱-۱- پیشینه پژوهش
۲۴	۲-۱- روند انجام کار
۲۵	۲- پیش نیازهای طرح
۲۵	۱-۲- پایگاه داده واژگان فارسی
۲۶	۲-۲- انتخاب نمونه‌های آزرشی
۲۷	۳-۲- دیگر پیش نیازها: علمی
۲۷	۳- استخراج ویژگی‌ها
۲۷	۱-۳- ویژگی‌های فیزیکی
۲۹	۲-۳- ویژگی‌های مفهومی
۳۰	۳-۳- ویژگی‌های آوایی
۳۲	۴-۳- کاهش ویژگی
۳۵	۴- آزمایش نمونه‌ها
۳۵	۱-۴- درخت تصمیم
۳۶	۲-۴- شبکه عصبی مصنوعی
۳۷	۵- نتیجه‌گیری
۳۹	مراجع
۴۱	پیوست ۱
۴۳	فصل ۳: تعیین شاعر به کمک روش‌های یادگیری ماشین
۴۴	۱- مقدمه
۴۵	۲- پیشینه پژوهش
۴۸	۳- استخراج ویژگی
۴۸	۱-۳- ویژگی‌های فیزیکی
۴۸	۲-۳- ویژگی‌های مفهومی
۴۸	۳-۳- ویژگی‌های آوایی
۴۹	۴- تعیین شاعر به کمک روش‌های دسته‌بندی
۴۹	۱-۴- درخت تصمیم
۵۰	۲-۴- شبکه عصبی مصنوعی
۵۱	۳-۴- تعیین شاعر توسط انسان
۵۳	۵- نتیجه‌گیری و پیشنهاد
۵۴	مراجع
۵۷	فصل ۴: تعیین ریشه زبانی واژگان فارسی و عربی به کمک شبکه عصبی مصنوعی
۵۸	۱- مقدمه
۵۹	۲- پیشینه پژوهش



۶۱	۳- شناخت مساله
۶۱	۱-۳- پیشینه زبان های جهان
۶۳	۲-۳- واژگان فارسی و لغات عربی
۶۴	۱-۲-۳- واژگان با ریشه ایرانی
۶۴	۲-۲-۳- واژگان با ریشه عربی
۶۴	۳-۲-۳- واژگان همانند
۶۵	۴- استخراج ویژگی
۶۷	۴- کد و گانه
۶۷	۲-۲- برآیند نفوذ
۶۹	۳-۴- کد دوگان ساده
۶۹	۴-۴- میانگین برآیند
۷۰	۵- تعیین ریشه زبان واژگان یک شبکه عصبی
۷۰	۱-۵- شبکه دوقطبی ورودی ساده
۷۰	۲-۵- شبکه دوقطبی با ورودی فشرده
۷۰	۳-۵- شبکه با ورودی فازی
۷۲	۴-۵- نتایج شبیه سازی و دسته بندی
۷۲	۶- نتیجه گیری
۷۲	مراجع

فصل ۵: تصحیح خودکار غلط های تایپی فارسی به کمک شبکه عصبی مصنوعی ترکیبی

۷۵	۱- مقدمه
۷۶	۲- پیشینه پژوهش
۷۹	۳- غلط تایپی
۸۰	۱-۳- روش های خودکار تصحیح غلط املائی
۸۱	۲-۳- غلط املائی در زبان فارسی
۸۲	۳-۳- غلط تایپی در زبان فارسی
۸۴	۴-۳- تصحیح املائی بر پایه فاصله
۸۶	۴- پیش نیازهای تصحیح خودکار غلط تایپی
۸۶	۱-۴- تعریف فاصله (شعاع همسایگی) کلیدها
۸۷	۲-۴- کدهای ورودی شبکه عصبی
۸۹	۳-۴- فرهنگ واژگان به کار گرفته شده در شبیه سازی
۸۹	۵- شبیه سازی
۸۹	۱-۵- شبیه سازی با شبکه عصبی هاپفیلد
۹۳	۲-۵- شبیه سازی با شبکه عصبی پرسپترون چندلایه
۹۵	۳-۵- ارزیابی نتایج
۹۵	۴-۵- بهبود مشکل ظرفیت شبکه
۱۰۰	۵- جمع بندی
۱۰۱	مراجع



۱۰۶	۲-۱- پیشینه خط میخی پارسی باستان
۱۰۶	۳-۱- ویژگی‌های خط میخی پارسی باستان
۱۰۸	۴-۱- مراحل یک روش بازشناسی نوری حروف
۱۰۸	۲- پیشینه پژوهش
۱۱۰	۳- گراف‌های فضایی و مکانی
۱۱۱	۴- بازشناسی نوری خط میخی
۱۱۱	۱-۴- ویژگی‌های خط میخی از دیدگاه پردازش تصویر
۱۱۲	۲-۴- پیش‌پردازش
۱۱۲	۳-۴- تشخیص ارای
۱۱۲	۴-۴- بازشناسی حروف
۱۱۲	۱-۴-۴- تشکیل گراف مکانی
۱۱۴	۲-۴-۴- کاهش گراف بانی
۱۱۵	۵- پیاده‌سازی و نتایج تجربی
۱۱۶	۶- نتیجه‌گیری و پیشنهاد
۱۱۷	مراجع

فصل ۷: تولید نام‌های زیبا و معنادار ایرانی به کمک الگوریتم ژنتیک با تابع برازندگی مبتنی بر شبکه عصبی مصنوعی

۱۲۱	۱- مقدمه
۱۲۲	۲- پیشینه پژوهش
۱۲۴	۳- ویژگی نام‌ها و واژگان ایرانی
۱۲۵	۱-۳- نام و معنای آن
۱۲۶	۲-۳- نام از نظر ساختاری
۱۲۷	۳-۳- ریشه زبانی نام‌های ایرانی
۱۲۷	۴-۳- زبان‌های آریایی
۱۲۸	۵-۳- زبان‌های ایرانی
۱۲۹	۶-۳- واژه‌سازی در زبان فارسی
۱۳۰	۷-۳- جابه‌جایی واژه‌ها در زبان فارسی
۱۳۱	۴- روش انجام کار
۱۳۲	۱-۴- مدل‌سازی
۱۳۴	۲-۴- ساخت پایگاه نام‌های آغازین
۱۳۵	۳-۴- تهیه پرسش‌نامه و نظرسنجی از جامعه آماری
۱۳۵	۴-۴- استخراج ویژگی
۱۴۰	۵-۴- شبکه عصبی به‌عنوان تابع برازندگی
۱۴۰	۶-۴- بردارهای ورودی و هدف
۱۴۴	۷-۴- الگوریتم ژنتیک
۱۴۵	۸-۴- تحلیل و ارزیابی
۱۴۷	۵- جمع‌بندی و کارهای آینده
۱۴۸	منابع

فصل ۸: پیش‌بینی نتایج بازی نرد به کمک روش‌های یادگیری ماشین

۱۵۳	۱- مقدمه
۱۵۴	



۱۵۵	۱-۱- تاریخچه
۱۵۵	۲-۱- تعاریف
۱۵۷	۳-۱- تفاوت‌های قوانین نرد ایرانی و دیگر کشورها
۱۵۷	۴-۱- راهبردهای نرد
۱۵۸	۲- پیشینه پژوهش
۱۵۹	۳- روش انجام کار
۱۵۹	۱-۳- کدگذاری
۱۶۰	۳- استنتاج ویژگی‌ها
۱۶۲	۳-۱- بهیمة پایگاه داده‌ها
۱۶۲	۴-۳- کاهش وابستگی ویژگی‌ها
۱۶۴	۳-۵- ابزار شبیه‌سازی
۱۶۴	۴- شبیه‌سازی و مدل‌سازی
۱۶۴	۱-۴- تحلیل آماری داده‌ها و الی
۱۶۵	۴-۲- شبکه عصبی مصنوعی
۱۶۵	۴-۳- دسته‌بندی بیزی ساده
۱۶۶	۴-۴- ارزیابی نتایج
۱۶۷	۵- جمع‌بندی
۱۶۸	منابع

فصل ۹: کد پرسک: روشی نو در کشف و تصحیح خطا

۱۷۱	۱- مقدمه
۱۷۲	۲- تعاریف
۱۷۵	۳- کد پرسک
۱۷۶	۱-۳- ساختار روش
۱۷۶	۲-۳- کدگذاری کد پرسک
۱۷۷	۳-۳- فرایند کشف خطا
۱۷۸	۴-۳- فرایند تصحیح خطا
۱۸۱	۳-۵- بیت‌های بازفرست
۱۸۱	۴- ویژگی‌های کد پرسک
۱۸۱	۱-۴- رابطه طول داده و سرباری
۱۸۲	۲-۴- شرایط رخداد حالت بازفرست
۱۸۴	۳-۴- احتمال رخداد حالت بازفرست
۱۸۵	۴-۴- شبیه‌سازی و نتایج
۱۸۷	۵- طراحی سخت‌افزاری
۱۸۷	۱-۵- مدار کدگذار
۱۸۵	۲-۵- مدار کدگشا
۱۸۹	۶- نتیجه‌گیری
۱۹۰	مراجع