

# داده کاوی با استفاده از تکنیکهای محاسبات نرم برای تشخیص کاراکترها

شهاب الدین شمشیربند , سعید ستایشی , محمود نقیب زاده

## چکیده

در این مقاله با استفاده از تکنیکهای محاسبات نرم ، روشی برای تشخیص کاراکتر بیان شده است.. ما بر مبنای تبدیل Hough هر الگوی پیکسلی از کاراکترها ، یک مجموعه ی فازی تعریف می کنیم. و مجموعه فازی اضافه شده را توسط یک T-norm تعیین می نماییم . ارتفاعهای این T-norm ها یک بردار ویژگی n-بعدی برای کاراکتر تشکیل می دهد. سپس یک بردار 3n-بعدی از بردار ویژگی n-بعدی بوسیله تعریف مجموعه فازی سه زبانی (یعنی ضعیف ، میانه رو و قوی ) برای هر عنصر ویژگی تولید می گردد .این بردار 3n-بعدی برای تمام الگوهای کاراکتر، ورودی برای شبکه عصبی چند لایه ای پرسپترون (Mlp) با یادگیری بوسیله پس انتشار خطاها، تشکیل می دهد. پس از انجام چند مرحله خروجی های چند لایه ای پرسپترون به ویژگی های ورودی بوسیله الگوریتم ژنتیک ، مجموعه ویژگی فازی ، انتخاب شده است. مزیت استفاده از توابع عضویت مجموعه فازی در خروجی mlp این است که فقط خروجی های با مقدار بالای عضویت ، در جستجو مورد استفاده قرار می گیرند. در طول این بازشناسی ، خروجی های با مقدار عضویت بالای مجموعه فازی برای جستجوی مبتنی بر لغتنامه بررسی می شوند که کاراکترهای مبهم را بر اساس دانش کلمه ای مشخص نمایند. این سیستم برای تشخیص کاراکترها درون اسناد انگلیسی چاپی پیاده سازی شده است.

کلمات کلیدی: محاسبات نرم , مجموعه های فازی , شبکه عصبی پرسپترون