به نام خدا



گزارش پروژه منطق فازی

درس هوش محاسباتی

دانشگاه اصفهان

دانشکده مهندسی کامپیوتر

استاد درس : دکتر حسین کارشناس

تهیه کنندگان :  
علیرضا دستمالچی ساعی، محمد حسین دهقانی

خرداد 1402

**مقدمه**

در این پروژه خواسته شده بود که از یک دیتاست مخصوص پیام­های ایمیل که مشخص می­کند ایمیل اسپم است یا نه، یک سیستم استنتاج فازی با یک مجموعه قانون بسازیم که با گرفتن ویژگی­های استخراج شده از متن یک ایمیل، اسپم بودن یک ایمیل را تشخیص دهیم.

**پیاده­سازی**

این برنامه به دنبال پیدا کردن بهترین مدل رده‌بندی اسپم و غیر اسپم برای داده‌های متنی است. برای این کار از الگوریتم ژنتیک استفاده شده است.

در این برنامه، ابتدا داده‌های متنی از یک فایل خوانده شده و سپس آن‌ها به یک دیتافریم تبدیل می‌شوند. سپس از روش TF-IDF برای استخراج ویژگی‌های مهم استفاده می‌شود و سپس با استفاده از روش کاهش بعد PCA، تعداد ویژگی‌ها (تعداد بعدها) کاهش داده می‌شود. در ادامه، با استفاده از روش feature\_selection تنها ویژگی­های مهم برای دسته­بندی انتخاب می­شوند.

**الگوریتم تکاملی**

برای الگوریتم تکاملی، از کتابخانه آماده Deap که یک کتابخانه قوی و جامع برای اجرای الگوریتم­های تکاملی است، استفاده کرده­ایم. ابتدا ژنوتایپ یا کروموزوم (Individual) و تابع ارزیابی (Evaluation) را برای toolbox کتابخانه تعریف می­کنیم.

نمایش ما (ژنوتایپ) به صورت یک دیکشنری می­باشد که در toolbox رجیستر شده و دارای 2 بخش می­باشد:

* قسمت Rules که شامل لیستی از قوانین است که به آن Rule base نیز می­گوین، بدین صورت که هر قانون X1 تا Xn و Y دارد. تعداد قوانین هم بین 50 تا 500 عدد می­باشد.
* قسمت Linguistic Functions که برای هر متغیر زبانی یک تابع درجه عضویت تعریف کرده­ایم که در ابتدا به صورت رندوم انتخاب می­شود.

سپس، برای تغییر، از توابعی که خودمان تعریف کرده­ایم استفاده می­کنیم:

* **بازترکیب:**

برای بازترکیب از عملگر تغییر، از one\_point\_crossover استفاده کرده­ایم.

* **جهش:**

برای جهش از عملگر تغییر، با یک احتمالی مقادیر قانون­هارا تغییر می­دهیم سپس با همان احتمال قبلی، توابع عضویت مخصوص هر مقدار زبانی، m و s را جهش می­دهیم.

برای تابع Evaluate که برازندگی هر ژنوتایپ (کروموزم) را حساب می­کند، با استفاده فرمول­هایی که در داک ذکر شده­اند، ساتفاده شده است.