



دانشگاه صنعتی امیرکبیر  
(پلی تکنیک تهران)  
دانشکده ریاضی و علوم کامپیوتر

ارائه دهنده:

عرفان نجفی

پروژه ششم

استاد درس:

دکتر مهدی قطعی

استاد دوم:

بهنام یوسفی مهر

## فهرست مطالب

۳.....	مقدمه
۴.....	توضیح درباره الگوریتم ژنتیک
۶.....	نحوه ارتباط الگوریتم ژنتیک به خوشه بندی
۸.....	توضیح مختصر کد

## مقدمه

ما یک دیتابیس داده داریم که شامل ۷ نوع داده است. می‌خواهیم استفاده الگوریتم ژنتیک را در مسائل خوشه بندی بررسی کنیم و پیاده‌سازی مدلی از آن را ببینیم. پس در ابتدا توضیحی از الگوریتم ژنتیک می‌دهیم و نحوه ارتباط آن را به این نوع مسائل بررسی می‌کنیم. بعد به پیاده‌سازی آن می‌پردازیم و در انتها به مقایسه آن با K-Means می‌پردازیم.

## توضیح درباره الگوریتم ژنتیک

الگوریتم‌های ژنتیک (Genetic Algorithms) بخشی از روش‌های محاسبات تکاملی هستند که الهام گرفته از فرآیندهای طبیعی تکامل زیستی می‌باشند. این الگوریتم‌ها برای حل مسائل بهینه‌سازی و جستجو مورد استفاده قرار می‌گیرند. اساس کار الگوریتم ژنتیک بر روی تولید، ارزیابی و انتخاب جمعیت‌ها استوار است تا به یک راه‌حل بهینه دست یابند.

### ساختار الگوریتم ژنتیک

الگوریتم ژنتیک شامل چندین بخش اصلی است:

اولین تعریف قبل از این بخش‌ها تعریف ژن است. تعریف کردن مفهوم یک راه‌حل به صورت ترجیحاً عددی که به آن کروموزوم می‌گوییم. هر کروموزوم شامل چندین ژن است که ژن در واقع بخش‌های خرد شده یک راه‌حل هستند.

#### ۱. جمعیت اولیه (Initialize population) :

در ابتدا، یک مجموعه‌ای از کروموزوم‌ها به صورت تصادفی یا بر اساس یک قاعده خاص تولید می‌شوند. این مجموعه، جمعیت اولیه را تشکیل می‌دهد.

#### ۳. تابع سنجش (Fitness Function) :

برای ارزیابی کیفیت هر کروموزوم، از یک تابع سنجش استفاده می‌شود. این تابع تعیین می‌کند که هر کروموزوم چقدر به هدف نزدیک است. کروموزوم‌هایی که برآزش بالاتری دارند، شانس بیشتری برای انتخاب و تولید مثل دارند. این تابع مانند یک هیوریستیک می‌ماند و میتواند میزان سرعت و کارایی برنامه را افزایش دهد.

#### ۴. انتخاب (Selection)

انتخاب کروموزوم‌ها برای تولید نسل بعدی بر اساس میزان کارایی (خرجی تابع سنجش) آن‌ها صورت می‌گیرد. روش‌های مختلفی برای انتخاب وجود دارد و چند تا از آن‌ها لیست میشود:

### ۱) انتخاب تناسبی (Roulette Wheel Selection):

این روش به این صورت است که هر کروموزوم به نسبت میزان برآزش خود شانس انتخاب شدن دارد، مشابه چرخ رولت که هر بخش از چرخ به نسبت احتمال خود انتخاب می‌شود.

### ۲) انتخاب تورنمنت (Tournament Selection):

در این روش، چندین کروموزوم به صورت تصادفی انتخاب و بهترین آن‌ها به عنوان والد برگزیده می‌شود.

### ۳) انتخاب رتبه‌بندی (Rank Selection):

در این روش، کروموزوم‌ها بر اساس ارزش خود رتبه‌بندی می‌شوند و شانس انتخاب شدن بر اساس رتبه آن‌ها تعیین می‌شود. (قدر مطلق نیست)

### ۵. تولید مثل (Crossover):

در این مرحله، دو کروموزوم والد با هم ترکیب می‌شوند تا کروموزوم‌های جدید (فرزندان) را ایجاد کنند. این فرآیند باعث ایجاد تنوع ژنتیکی در جمعیت و امکان بهبود راه‌حل‌ها می‌شود. این ترکیب می‌تواند به صورت تک نقطه‌ای، دو نقطه‌ای یا چند نقطه‌ای باشد. برای مثال، در روش تک نقطه‌ای، یک نقطه تصادفی در کروموزوم انتخاب شده و قسمت‌های قبل و بعد از آن نقطه بین دو والد تبادل می‌شود، روش دونقطه‌ای اینگونه است که دو نقطه به تصادف انتخاب شده و فضای بین آن‌ها از یک والد و بخش‌های خارجی از والد دیگر. همین روش را میتوان تا  $n$  نقطه ای گسترش داد.

### ۶. جهش (Mutation):

در این مرحله، به صورت تصادفی برخی از بیت‌ها یا ژن‌های کروموزوم تغییر می‌کنند. تغییر هر کدام از ژن‌ها با یک احتمال یکسان صورت می‌گیرد. این تغییرات به منظور ایجاد تنوع و جلوگیری از به دام افتادن الگوریتم در بهینه‌های محلی انجام می‌شود. برای مثال، در کروموزوم دودویی، یک بیت ۰ به ۱ یا برعکس تبدیل می‌شود.

الگوریتم ژنتیک الهام گرفته از فرآیند تکامل طبیعی است. در طبیعت، گونه‌ها بر اساس ژنتیک خود و محیط زندگی‌شان تکامل می‌یابند. در انتخاب طبیعی، موجودات با ویژگی‌های بیشتر برای زنده ماندن و تولید مثل شانس بیشتری دارند. این ویژگی‌ها از طریق ژن‌های والدین به نسل‌های بعدی منتقل می‌شود.

الگوریتم‌های ژنتیک با الهام از طبیعت، یکی از موثرترین روش‌های بهینه‌سازی و جستجو در هوش مصنوعی هستند. این الگوریتم‌ها با استفاده از فرآیندهای انتخاب، تولید مثل و جهش، توانایی پیدا کردن راه‌حل‌های بهینه را دارند و در مسائل مختلفی از جمله بهینه‌سازی توابع پیچیده، طراحی مهندسی و یادگیری ماشین کاربرد دارند.

## نحوه ارتباط الگوریتم ژنتیک با خوشه‌بندی چگونه است؟

الگوریتم‌های ژنتیک می‌توانند به عنوان یک روش بهینه‌سازی برای خوشه‌بندی داده‌ها مورد استفاده قرار گیرند. این الگوریتم‌ها با استفاده از فرآیندهای تکاملی به یافتن خوشه‌های بهینه کمک می‌کنند و می‌توانند برای مسائل پیچیده و با فضای جستجوی بزرگ بسیار موثر باشند.

با مشخص کردن هدف در یک دیتاست ما می‌توانیم با الگوریتم ژنتیک جواب‌ها نزدیک به هدف یا خود هدف را پیدا کنیم و به عنوان خوشه معرفی کنیم.

## توضیح مختصر درباره پیاده سازی

