

به نام خرد

پروژه سوم هوش مصنوعی

جستجوی محلی

بهار ۱۴۰۴

بخش یکم: گزینش ویژگی با الگوریتم ژنتیک

یکی از چالش های حوزه یادگیری ماشین، حجم و ابعاد بالای داده های آموزشی است که سبب کند شدن فرآیند یادگیری مدل و همچنین نیازمند توان محاسباتی بالا است. یکی از روش های مقابله با این چالش گزینش ویژگی (Feature Selection) می باشد که در آن به دنبال بهترین ویژگی ها که همبستگی بالایی با ویژگی هدف دارند هستیم و سایر ویژگی های نا مرتبط را از مجموعه داده ها حذف می کنیم. به نمونه زیر توجه کنید.

فرض کنید می خواهیم یک مدل یادگیری ماشین برای پیش بینی ابتلای افراد به دیابت بر اساس ویژگی های بیمار طراحی کنیم. ویژگی هایی که از بیماران در اختیار داریم شامل موارد زیر هستند.

۱. قد ۲. وزن ۳. سن ۴. شاخص توده بدن ۵. گروه خونی ۶. گلوکز خون

از میان شش ویژگی بالا شاخص توده بدن (BMI) خود از قد و وزن بیمار ساخته می شود در نتیجه می توان دو ویژگی خام قد و وزن را حذف کرده و ویژگی ترکیبی توده بدن را نگه داشت.

همچنین میان احتمال ابتلا به دیابت و سن بیمار و همچنین سطح گلوکز خون همبستگی بالایی وجود دارد و این دو ویژگی در یادگیری مدل بسیار موثرند.

اما میان ویژگی گروه خونی بیمار و دیابت چندان ارتباط معنا داری وجود ندارد و بیشتر سبب گیج شدن مدل و افزایش حجم محاسبات می شود.

در نتیجه از میان شش ویژگی موجود در پایگاه داده بیمارستان سه ویژگی ۳. سن ۴. شاخص توده بدن ۶. گلوکز خون را برای آموزش مدل هوش مصنوعی خود بر می گزینیم.

در بسیاری از مسائل دنیای واقعی تعداد ویژگی ها بسیار بالا است و امکان استدلال برای گزینش بهترین گروه ویژگی ها همواره ممکن نیست. الگوریتم ژنتیک می تواند برای پیدا کردن بهترین زیر مجموعه از کل مجموعه ویژگی ها به ما کمک کند.

نشانه گذاری

هر ژن را می توان یک بیت در نظر گرفت که در آن 0 به معنای حذف ویژگی و 1 به معنای حفظ ویژگی است و در هر کروموزم به تعداد ویژگی های دادگان ژن قرار می دهیم.

به طور مثال برای مسئله بالا رشته کروموزمی که در آن سه ویژگی ۳. سن ۴. شاخص توده بدن ۶. گلوکز خون بر گزیده شدند رشته ای به طول ۶ و به شکل $[0, 0, 1, 1, 0, 1]$ می باشد.

تابع برازش

برای تعیین برازش (Fitness) یک کروموزم می توان با ویژگی های برگزیده یک مدل را آموزش داد و دقت عملکرد به دست آمده را به عنوان برازش در نظر گرفت. همچنین برای جلوگیری از گزینش همه ویژگی ها برای کروموزم هایی که ژن های یک زیادی دارند جریمه در نظر گرفت.

پروژه

فرض کنید یک مجموعه داده ناشناخته در اختیار دارید که دارای ۱۳ ویژگی است که هیچ اطلاعاتی در باره ویژگی ها نداریم و می خواهیم ۵ ویژگی که دارای بیشترین همبستگی با ویژگی هدف هستند را پیدا کنیم.

مجموعه داده ناشناخته با نام `unknown.data` به همراه کد پایتون پیش پردازش دادگان و پیاده سازی تابع برازش با نام `fitness.ipynb` پیوست شده است. برای درک بهتر، سلول آخر کد را چندین بار اجرا کنید و خروجی را مشاهده کنید.

فاز های گزینش، آمیزش و جهش الگوریتم ژنتیک را پیاده سازی کنید. سپس با تنظیم ابر پارامتر های تعداد جمعیت، نسل و نرخ جهش، الگوریتم را اجرا کرده و کروموزم با بیشترین برازش را بیابید.

بخش دوم: بهینه سازی با الگوریتم سرد سازی گام به گام

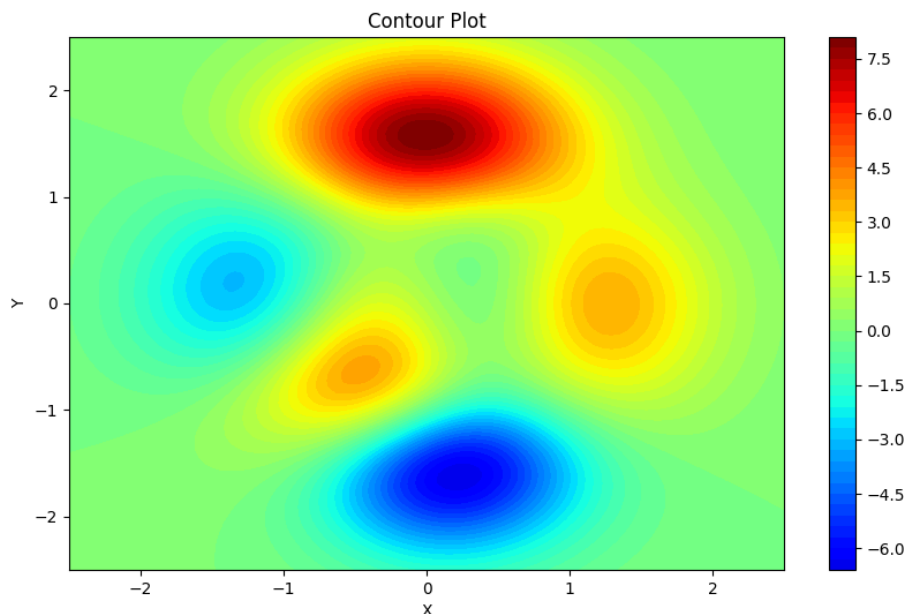
برای یافتن کمینه یا بیشینه یک تابع ریاضی می توان از الگوریتم سرد سازی گام به گام بهره برد.

الگوریتم Simulated Annealing یا به اختصار SA یک روش جستجوی تصادفی هوشمند است که از یک نقطه تصادفی آغاز شده و به تصادف یک جهت را برای حرکت بعدی بر می گزیند که با گذر زمان و کاهش تدریجی دما، تمرکز الگوریتم از پذیرش حرکات بدی که تابع برازش را می کاهش دهد (برای فرار از بهینه های محلی) به سمت پذیرش حرکات خوبی که فقط تابع برازش را می افزایش دهد معطوف می شود. الگوریتم SA را پیاده سازی کرده و با تنظیم ابر پارامتر های دما، نرخ کاهش دما و تعداد تکرار الگوریتم به پرسش های صفحه پایانی در باره توابع زیر پاسخ دهید.

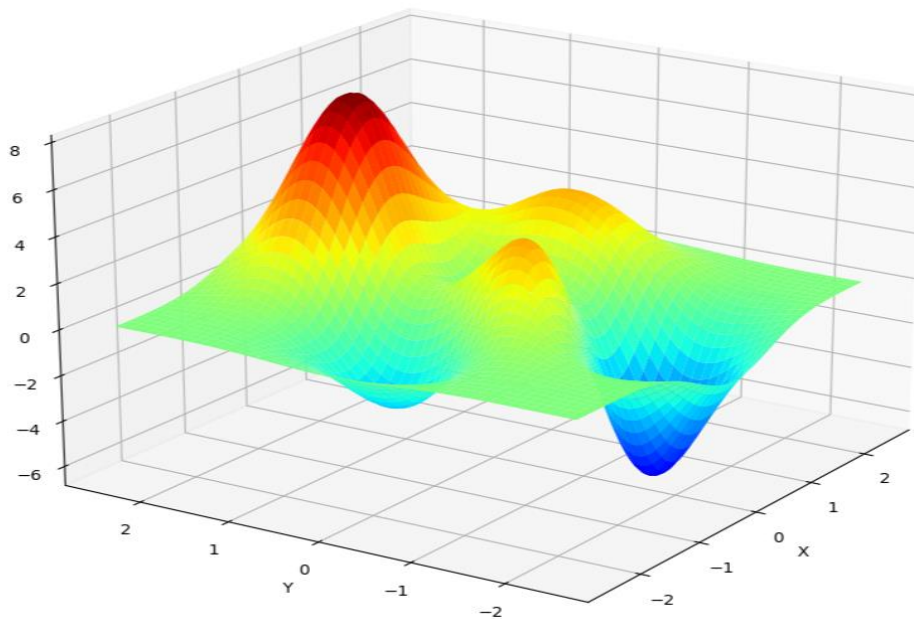
• ضابطه تابع یکم

$$f(x, y) = 3(1 - x)^2 e^{-x^2 - (y+1)^2} - 10 \left(\frac{x}{5} - x^3 - y^5 \right) e^{-x^2 - y^2} - \frac{1}{3} e^{-(x+1)^2 - y^2}$$

• نمودار کانتور تابع (دو بعدی)



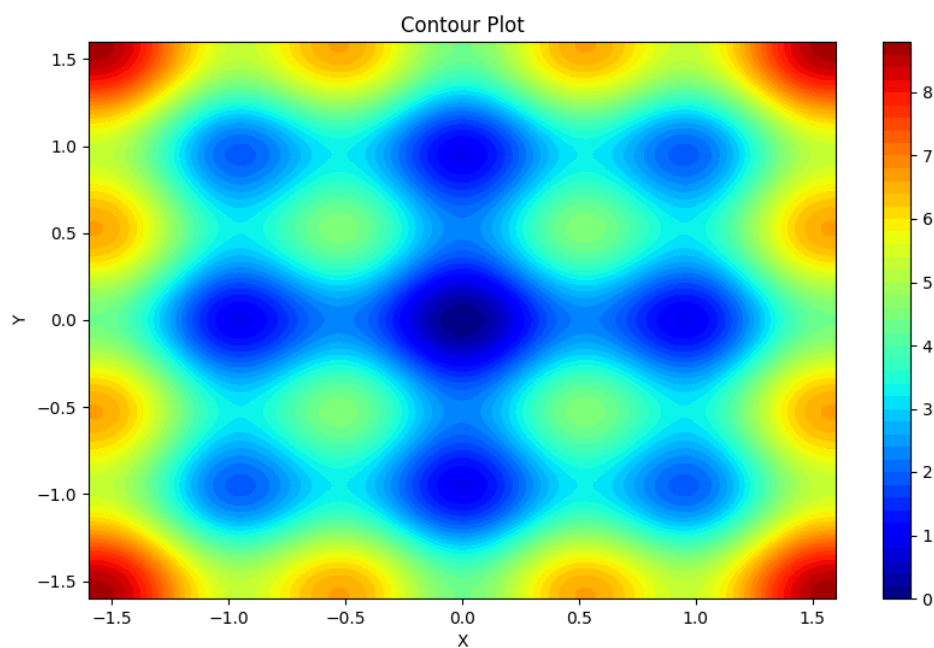
- نمودار سه بعدی تابع



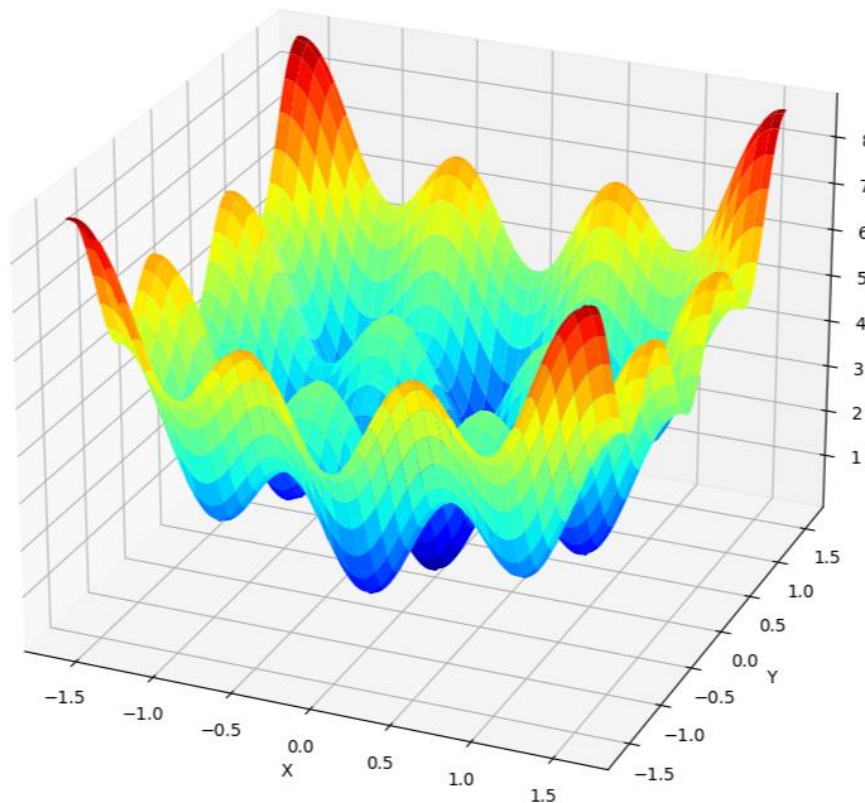
- ضابطه تابع دوم به نام Rastrigin Function برای دو متغیر

$$f(x, y) = 20 + (x^2 - 10 \cos(2\pi x)) + (y^2 - 10 \cos(2\pi y))$$

- نمودار کانتور تابع (دو بعدی)



- نمودار سه بعدی تابع



پرسش

برای درک بهتر توابع دو فایل با نام های `objective.py` و `rastrigin.py` پیوست شده است. آن ها را اجرا کنید و نمودار تعاملی سه بعدی توابع را از زوایای مختلف مشاهده کنید. همچنین ضابطه توابع را نیز برای پیاده سازی خود کپی کنید.

۱. مختصات و مقدار کمینه جهانی و بیشینه جهانی تابع یکم را بیابید.

محدوده جستجو $[-3, +3]$ برای هر دو متغیر x, y

۲. مختصات و مقدار کمینه جهانی تابع دوم Rastrigin بیابید.

محدوده جستجو $[-1.5, +1.5]$ برای هر دو متغیر x, y

توجه کنید این تابع پر از کمینه های محلی ست. در نتیجه ابر پارامتر ها را به خوبی تنظیم کنید و

در صورت نیاز چندین بار الگوریتم را اجرا کنید.

نکته: برای یافتن کمینه کافی ست تابع برازش را قرینه کنید.