



**دانشگاه صنعتی امیرکبیر**  
(پلی تکنیک تهران)

دانشکده ریاضی و علوم کامپیوتر

استاد درس: دکتر مهدی قطعی

استاد کارگاه: بهنام یوسفی مهر

پاییز ۱۴۰۲

**راهنمای تمرین ۷**

درس هوش مصنوعی و کارگاه



در این تمرین باید مقاله‌ی **Swarm intelligence for self-organized clustering** که از **این لینک** قابل دریافت است را مطالعه و پیاده‌سازی کرده و نتایج پیاده‌سازی خود را بیان کنید. همچنین **کتاب نویسنده‌ی مقاله** هم در اختیار شما قرار خواهد گرفت. مطالب مقاله را می‌توانید با جزئیات بیشتر از روی کتاب هم دنبال کنید. برای آشنایی با نمادگذاری‌های نویسنده توصیه می‌شود که فصل‌های ابتدایی کتاب را هم مطالعه کنید.

## مطالعه و درک مقاله

در این بخش پرسش‌هایی مطرح شده است که باید در گزارش خود به طور مستقیم یا غیرمستقیم به آنها پاسخ دهید. این پرسش‌ها ذهنیتی از پروژه برای شما ایجاد خواهد کرد و لذا قبل از مطالعه‌ی مقاله توصیه می‌شود که همگی آن‌ها را مطالعه کنید. تمامی پرسش‌ها طوری انتخاب شده‌اند که پاسخ به آنها در درک بهتر پروژه و پیاده‌سازی آن به شما کمک کند.

- هدف DBS چیست و در چه دسته‌ای از روش‌ها قرار می‌گیرد؟
- تعریفی از Pswarm ارائه کرده و سه ایده‌ای که در این مفهوم استفاده می‌شود را ذکر کنید. همچنین درک خود را به طور مختصر از این سه ایده بیان کنید.
- چگونه Pswarm یک گام را مدل می‌کند و چه تضمینی در هر گام وجود دارد؟
- فضای خروجی را توصیف کرده، مشکلات مرزی فضا را بیان کنید و بگویید چگونه بر این مشکلات فائق می‌شویم.
- درمورد ابعاد فضای مشبک<sup>۱</sup> بحث کنید و تفاوت آن با مقالات دیگر را بیان کنید. در این مورد با چه چالش‌هایی روبه‌رو هستید؟ آیا نویسنده به طور دقیق درمورد ابعاد فضا توضیح داده است؟
- چگونه فضای مشبک را در حالت‌های مختلف تولید می‌کنید؟ علت انتخاب پارامترها را با توجه به توضیحات نویسنده توجیه کنید.
- ارتباط scent با payoff را بیان کنید.
- تمام تبدیلاتی که باید از مختصات کارترزین به قطبی و از قطبی به کارترزین انجام دهید را با اشاره به پیاده‌سازی خود بیان کنید.

---

<sup>۱</sup>grid

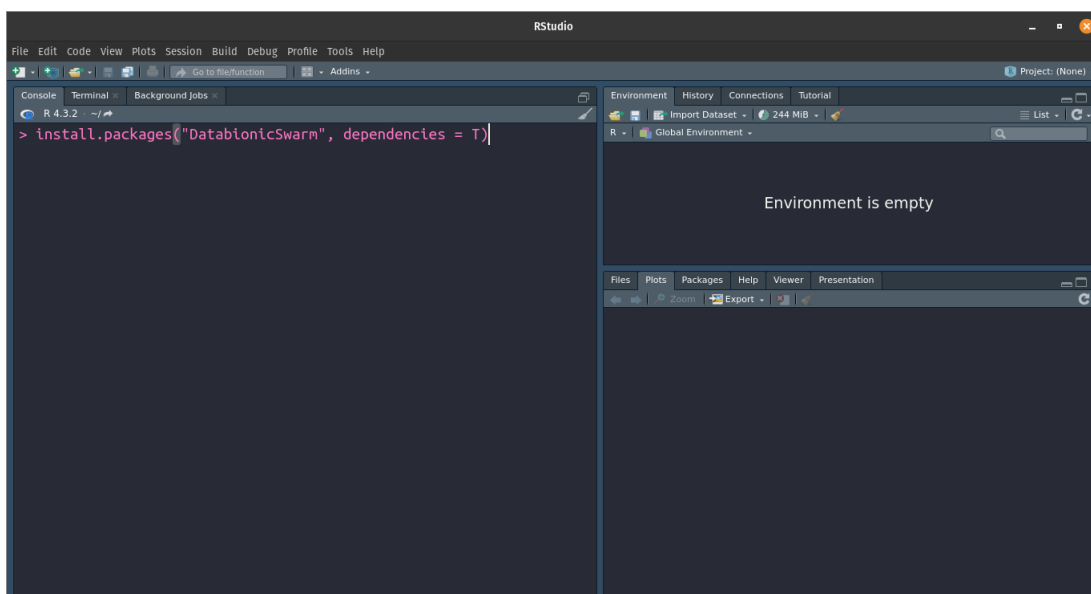


## اجرای برنامه در محیط R Studio

برای آشنایی بهتر شما با مراحل پیاده‌سازی، برنامه را یک بار در محیط R Studio اجرا کرده و اجرای آن را در مراحل مختلف بررسی می‌کنیم.

### گام صفر: نصب برنامه در R

علاوه بر خود برنامه با دستور زیر تمام وابستگی‌های برنامه را هم در R نصب می‌کنیم. به همین دلیل ممکن است این قسمت زمان طولانی‌تری نیاز داشته باشد.



### گام اول: تصویر کردن داده‌های با ابعاد بالا

در این گام یک تصویر دو بعدی از یک ماتریس ورودی ایجاد می‌کنیم. ماتریس ورودی یک ماتریس  $n \times n$  بوده که  $n$  تعداد رکوردها است. دستوراتی که در کنسول وارد می‌کنیم در ادامه نشان داده شده است.



```
Console Terminal Background Jobs
R 4.3.2 ~ /
> library(DatabionicSwarm)
> library(GeneralizedUmatrix)
> data('Hepta')
> View(Hepta)
> InputDistances = as.matrix(dist(Hepta$Data))
> View(InputDistances)
> projection = Pswarm(InputDistances)
|=====| 100%
```

در اینجا پس از فراخوانی کتابخانه‌های مورد نیاز، از داده‌های **Hepta** استفاده کرده‌ایم. **Hepta** ۲۱۲ رکورد دارد و لذا پس از ایجاد ماتریس فاصله با کمک تابع فاصله‌ی **dist**، یک ماتریس متقارن  $212 \times 212$  به دست می‌آید که خروجی آن را در شکل زیر مشاهده می‌کنید.

	1	2	3	4	5	6	7	8	9
1	0.00000000	0.08058895	0.04781043	0.09214454	0.03886835	0.09699465	0.09783282	0.06061589	0.10267481
2	0.08058895	0.00000000	0.08453525	0.06323338	0.10074380	0.04796023	0.13600445	0.09067025	0.15028515
3	0.04781043	0.08453525	0.00000000	0.07410562	0.05199453	0.11655407	0.13884176	0.08300249	0.09621151
4	0.09214454	0.06323338	0.07410562	0.00000000	0.08676341	0.08379470	0.14497584	0.07903343	0.10998261
5	0.03886835	0.10074380	0.05199453	0.08676341	0.00000000	0.11133235	0.09440673	0.04455874	0.06398915
6	0.09699465	0.04796023	0.11655407	0.08379470	0.11133235	0.00000000	0.11084723	0.08240419	0.15640072
7	0.09783282	0.13600445	0.13884176	0.14497584	0.09440673	0.11084723	0.00000000	0.06914267	0.12598618
8	0.06061589	0.09067025	0.08300249	0.07903343	0.04455874	0.08240419	0.06914267	0.00000000	0.07935143
9	0.10267481	0.15028515	0.09621151	0.10998261	0.06398915	0.15640072	0.12598618	0.07935143	0.00000000
10	0.08356635	0.08882601	0.05254551	0.03924898	0.06832553	0.11335779	0.14777159	0.07931329	0.08222241
11	0.12395455	0.10368294	0.10325634	0.04212453	0.10652446	0.11405738	0.16166011	0.09559737	0.10432641

Showing 1 to 11 of 212 entries, 212 total columns

همچنین خروجی الگوریتم **Pswarm** به صورت زیر است. توجه کنید که ابعاد داده‌های ورودی که برای آن ماتریس فاصله محاسبه شد در ابتدا ۳ بود و پس از فراخوانی الگوریتم **Pswarm** به ۲ تبدیل شد.

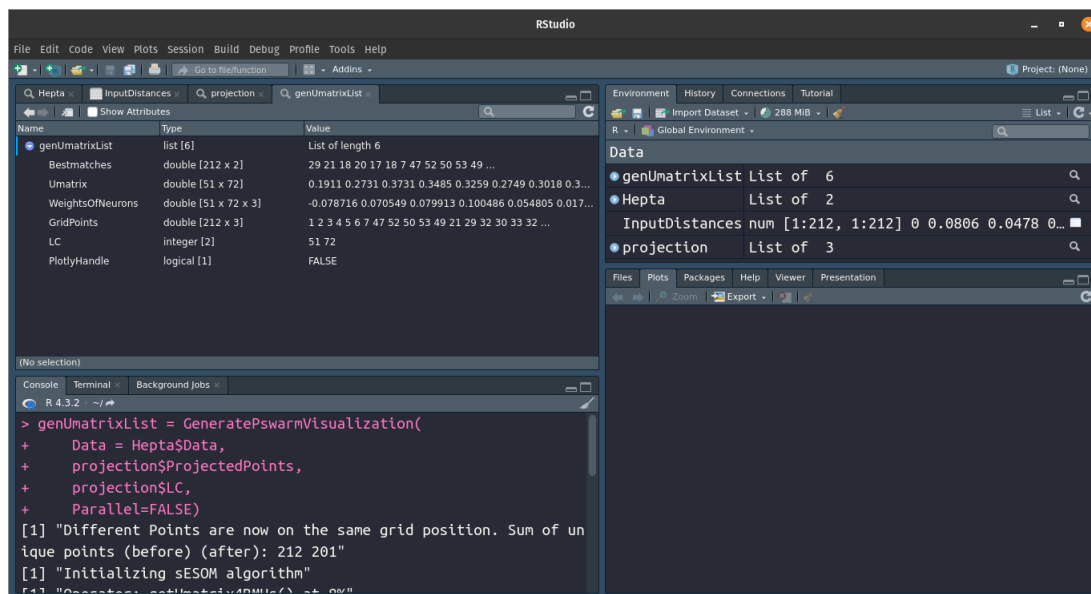
Name	Type	Value
projection	list [3]	List of length 3
ProjectedPoints	double [212 x 2]	6.5 45.5 51.0 49.0 51.5 48.0 19.0 27.0 30.0 28.0 31.0 30...
LC	double [2]	51 72
Lines	double [1]	51
Columns	double [1]	72
Control	list [3]	List of length 3
stressverlauf	double [1244]	6.76 12.42 16.98 23.10 30.46 35.22 ...
epocheadiusred...	integer [18]	129 6 50 4 4 42 ...
LetzteEpocheStr...	double [1]	Inf

projection



## گام دوم: محاسبه‌ی Generalized U-Matrix

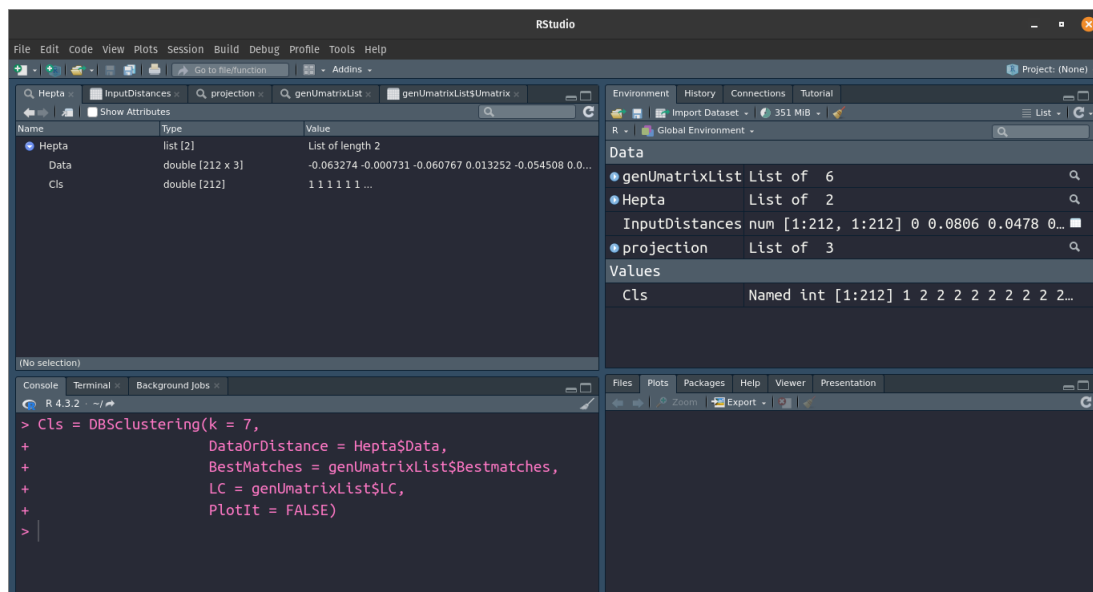
در گام بعدی به محاسبه کردن Generalized U-Matrix می‌پردازیم.



همان طور که در شکل بالا مشاهده می‌شود خود داده‌ها، تصویر دوبعدی داده‌ها و اندازه‌ی فضا ورودی‌های مهمی هستند که به تابع داده می‌شوند. خروجی مهم‌تر نیز ماتریس UMatrix با ابعاد  $51 \times 72$  است.

## گام سوم: خوشه‌بندی

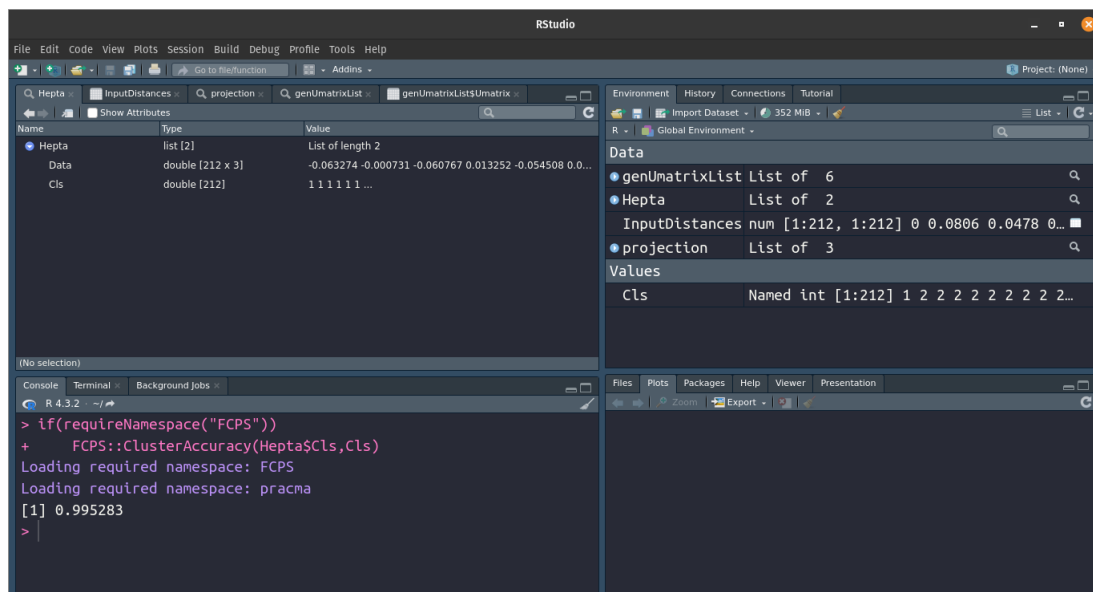
در این گام عملیات خوشه‌بندی را انجام می‌دهیم.



همان طور که مشاهده می‌شود تعداد  $k = 7$  خوشه را به عنوان پارامتر به تابع داده‌ایم و در خروجی برچسب هر رکورد که مربوط به یک خوشه است را دریافت کرده‌ایم.

## گام چهارم: دقت خوشه‌بندی

محاسبه‌ی دقت خوشه‌بندی در R در شکل زیر نشان داده شده است.



مشاهده می‌کنیم که خوشه‌بندی روی Hepta دقت بسیار بالایی داشته است.

## پیاده‌سازی

### گام اول: تصویر کردن داده‌های با ابعاد بالا

ورودی این گام یک ماتریس فاصله <sup>۲</sup> است. در این قسمت باید یک ماتریس فاصله از روی داده‌های ورودی تشکیل دهید و آن را به عنوان ورودی به الگوریتم Pswarm بدهید. خروجی اصلی این قسمت یک ماتریس  $n \times 2$  خواهد بود که  $n$  تعداد رکوردهاست.

### گام دوم: محاسبه‌ی Generalized U-Matrix

در این گام باید خود داده‌ها، تصویر دوبعدی داده‌ها (ماتریس  $n \times 2$ ) و ابعاد فضا را به عنوان ورودی به برنامه بدهید و در خروجی یک U-Matrix از برنامه دریافت کنید. در پیاده‌سازی این قسمت می‌توانید از [این مقاله](#) استفاده کنید. در این گام نیاز به تشکیل نقشه‌ی توپوگرافیک <sup>۳</sup> نیست.

Distance Matrix<sup>۲</sup>  
Topographic Map<sup>۳</sup>



### گام سوم: خوشه‌بندی

برای آشنایی با خوشه‌بندی در DBS می‌توانید صفحه‌ی ۱۰۴ کتابی که در اختیار شما قرار داده شده است را مطالعه کنید. خروجی این قسمت یک بردار  $n$  تایی شامل برچسب‌های متناظر هر رکورد است. پارامتر تعداد خوشه‌ها را با استفاده از اطلاعات قبلی روی دیتاستی که از آن استفاده می‌کنید به برنامه بدهید. در این بخش می‌توانید خودتان قسمت‌هایی را ساده‌سازی کنید، مهم آن است که آنچه را متوجه شده‌اید تا حد توان پیاده‌سازی کنید.

### گام چهارم: دقت خوشه‌بندی

رابطه‌ای که برای به دست آوردن دقت خوشه‌بندی استفاده می‌کنیم ساده است:

$$Accuracy = \frac{\# of True Positives}{\# of Records}$$

اما توجه داشته باشید که مسئله‌ی ما خوشه‌بندی است و نه طبقه‌بندی. پس باید از توابع مخصوصی استفاده کنیم.

### معیارهای سنجش تمرین

- پیاده‌سازی باید با زبان پایتون انجام شود.
- در گزارش خود باید به پرسش‌هایی که در بخش مطالعه و درک مقاله مطرح شدند به طور مستقیم یا غیرمستقیم پاسخ دهید.
- استفاده از کدها و مراجع دیگر با ذکر منبع بلامانع است.
- نحوه‌ی پیاده‌سازی خود را در گزارش شرح داده و استفاده از پارامترهای مختلف را با ارجاع به مقالات و یا کتابی که در اختیار شما قرار داده شده است توجیه کنید.
- استفاده از هر دیتاستی برای نتیجه‌گیری بلامانع است. همچنین می‌توانید از همان دیتاست‌های داخل مقاله هم استفاده کنید.
- هر گونه ساده‌سازی را در گزارش خود بیان کنید.

### نحوه‌ی پیوست کد به گزارش

اگر کدتان را در Google Colab پیاده‌سازی کرده‌اید لینک آن را در انتهای گزارش اضافه کنید. همچنین دسترسی به کد از طریق لینک را نیز در فایل Colab خود فعال کنید. می‌توانید کد را در Github بارگذاری کرده و لینک Github را به انتهای گزارش اضافه نمایید.