

این جلسه با گزارشی از اجرای الگوریتم pagerank بر روی مجموعه داده ها آغاز شد. در ادامه با اشاره به اجرای این الگوریتم در شرایط جهت دار یا بدون جهت بودن گراف نتایج زیر ارائه گردید:

```
>>> V_dir.join(V_undir, V_dir.id == V_undir.id).filter(V_dir.pagerank != V_undir.pagerank).select(V_dir.pagerank.alias('directed pagerank'), V_undir.pagerank.alias('undirected pagerank'), V_dir.id, V_dir.type).show()
```

directed pagerank	undirected pagerank	id	type
1.7245591080167213	1.4464517554276748	0439394422	p
0.6691289339104879	0.41924393256562453	0763030945	p
1.255479030636173	1.0442782608730694	3866811713	p
0.5518589145653507	1.80293772228192	A02836981FYG9912C66F	u
0.5518589145653507	0.432565966075097	A1016H15N3HBYY	u
0.5518589145653507	1.2429450246712963	A101AXLTZA3BT1	u
0.5518589145653507	0.37669143062264604	A101PGIDQR898L	u
0.5518589145653507	0.48500378379969716	A101ZIBWCYK7N	u
0.5518589145653507	0.39394651991643165	A102A7JAILVBB	u
0.5518589145653507	0.3919667077118938	A102ZH210YLV45	u
0.5518589145653507	0.3773461479223876	A10369XZ10GHJ2	u
0.5518589145653507	0.4481719448556828	A10388IPSD4E01	u
0.5518589145653507	0.389803175545804	A103WYQ70SZ9H0	u
0.5518589145653507	0.36693554544275403	A1043VSDGKG8UU	u
0.5518589145653507	0.4110597112424788	A1053F7LH50WL3	u
0.5518589145653507	0.45043640057686607	A106KQMU8Y9SDU	u
0.5518589145653507	0.41247939940982675	A108KKY6PMCEOT	u
0.5518589145653507	0.6662362809438488	A1090LTJNITNXT	u
0.5518589145653507	0.9355801632219716	A109J7F0H0K819	u
0.5518589145653507	0.8455080710226474	A10AGXV1HU4ET2	u

only showing top 20 rows

شکل ۱: نتایج اجرای الگوریتم pagerank با فرض جهت دار یا بدون جهت بودن گراف (id شناسه گره در شبکه و type نوع گره است که ۲ حالت p به معنی محصول و u به معنی کاربر دارد).

همانطور که در شکل ۱ مشخص است مقدار محاسبه شده pagerank در دو حالت متفاوت است.

```
>>> dirV.select(dirV.id, dirV.type, dirV.pagerank).orderBy(dirV.pagerank).show()
+-----+-----+-----+-----+-----+-----+
| id      | type | pagerank |
+-----+-----+-----+
| A14QSWQT2S607F | u | 0.5518589145653507 |
| A170N8URODGXW | u | 0.5518589145653507 |
| A1A3WX18ZR2075 | u | 0.5518589145653507 |
| A1ANRQS0VR58DP | u | 0.5518589145653507 |
| A1ENN8X0FHTHNG | u | 0.5518589145653507 |
| A1E0TMK87GQ70C | u | 0.5518589145653507 |
| A1FNAJTB1G0Z85 | u | 0.5518589145653507 |
| A1LIKAGYMIGWES | u | 0.5518589145653507 |
| A1MND25B9H9S1S | u | 0.5518589145653507 |
| A1OAU966J6088 | u | 0.5518589145653507 |
| A1QAFQHULB2FE3 | u | 0.5518589145653507 |
| A1T5U09QJRP5SY | u | 0.5518589145653507 |
| A1U8GA140DBET | u | 0.5518589145653507 |
| A23AFD1LEGO55T | u | 0.5518589145653507 |
| A23CQZDODYXNR0 | u | 0.5518589145653507 |
| A26EFIECCCTCVZ | u | 0.5518589145653507 |
| A2B0L3LTSCHB8F | u | 0.5518589145653507 |
| A2D1Y38202F094 | u | 0.5518589145653507 |
| A2P49HIFIH0UWT | u | 0.5518589145653507 |
| A2P5IJ685EMI03 | u | 0.5518589145653507 |
+-----+-----+-----+
only showing top 20 rows
```

شکل ۲: در حالتی که گراف جهت دار در نظر گرفته شود واضح است که مقدار pagerank برای تمامی گره های کاربر مقدار برابر خواهد داشت.

پس از بیان اینکه ماهیت گراف مورد نظر بدون جهت است و الگوریتم pagerank باید در شرایط بدون جهت بودن گراف اجرا شود ویژگی دویخشی^۱ بودن گراف مطرح شد. با توجه به اینکه گره های موجود در گراف به ۲ نوع کاربر یا محصول تقسیم می شوند و یال های گراف فقط بین این ۲ گروه قرار دارند و هیچ یالی بین خود این گروه ها وجود ندارد پس می توان نتیجه گرفت که این گراف، یک گراف دویخشی است. با بیان این موضوع سوالات زیر مطرح شد.

آیا سرعت و نحوه اجرای الگوریتم pagerank زمانی که گراف دویخشی باشد با حالت عادی متفاوت است؟

آیا راهی برای سریع تر کردن اجرای این الگوریتم و بهینه کردن آن در گراف دویخشی وجود دارد؟

برای یافتن پاسخ این سوالات دو مقاله زیر به صورت کلی مورد بررسی قرار گرفتند:

BiRank: Towards Ranking on Bipartite Graphs^۲

PAGERANK FOR NETWORKS, GRAPHS AND MARKOV CHAINS^۳

در آخر درمورد بخش های مختلف مجموعه داده ها بحث انجام گرفت و قرار شد در حال حاضر کار ها بر روی بخش videogames مجموعه داده های آمازون دنبال شود و در آینده در صورت تمایل ۲ بخش مرتبط با videogames اضافه شود.

به عنوان سخن پایانی قرار شد جلسه آینده پیاده سازی های اولیه الگوریتم pagerank در شرایطی که گراف بدون جهت و دویخشی است انجام گیرد و مطالعاتی درمورد پیاده سازی الگوریتم hobs and authority انجام شود.

1 Biprtite

2 [Xiangnan He, Ming Gao Member, IEEE, Min-Yen Kan Member, IEEE and Dingxian Wang; 2017](#)

3 C. ENGSTRÖM, S. SILVESTROV; 2017