
Analyzing Elective Choices of Bachelor Students at Innopolis University using Basket Market Analysis

— Magzumov Davlatkhodzha —
B20CS01

Introduction

In my internship, I aimed to explore the connection between Bachelor students' academic tracks and their elective choices at Innopolis University. To achieve this, I employed Basket Market Analysis, a data-driven approach widely used in market research and recommendation systems.

Introduction

Basket Market Analysis allowed us to identify patterns and associations within the elective choices made by students, offering valuable insights into their decision-making processes. By understanding these preferences, Innopolis University can better tailor its elective offerings to enhance students' academic journeys.

Methodology

Data was collected from student distributions in tracks and elective course distribution. Its structure was modified to make it anonymized and convenient for the experiment.

At all, there was chosen a 2 year span of tech and humanitarian courses for BS21 students.

Methology

Track		Tech22		Hum22	Tech23		Hum23
0	AI	Front-end Web Development	Introduction to Career Development for IT-spec...	Advanced Statistical and Econometric Methods			Design thinking
1	AI	Computer Graphics in Game Development		Career and leadership	Advanced Statistical and Econometric Methods		Design thinking
2	AI	Front-end Web Development		Tech Startup Design	Data Wrangling and Visualization		Social Entrepreneurship
3	AI	Introduction to Sensing and Data Acquisition	Developing Entrepreneurial Skills for a Startu...		Applied Machine Learning		Basics of Product Management
4	AI	Cross-platform Mobile Development with Flutter	Personal Efficiency Skills of IT-specialist	Advanced Statistical and Econometric Methods			Startup CEO Toolkit
...
185	SD	Computer Graphics in Game Development	Introduction to Career Development for IT-spec...		Front-end Web Development	Introduction to Public Speaking for IT-specialist	
186	SD	Functional Programming in Haskell	Personal Efficiency Skills of IT-specialist		System Programming	Introduction to Public Speaking for IT-specialist	
187	SD	Front-end Web Development	Introduction to Career Development for IT-spec...		Programming in Python		Design thinking
188	SD	Lambda-Calculus, Algebra, Machinery and Logic ...	Introduction to Career Development for IT-spec...	Challenges of Object Oriented Programming			Social Entrepreneurship
189	SD	Functional Programming in Haskell		Career and leadership	Applied Machine Learning	Personal Efficiency Skills of IT-specialist	

190 rows × 5 columns

Pic. 1 Dataframe for pandas

Methodology

Once the data was collected, I proceeded with the application of Basket Market Analysis (MBA) to explore the relationships between academic tracks and elective choices. MBA involves identifying frequent itemsets and association rules to uncover patterns within the data.

To conduct the analysis, we transformed the survey data into transactional data, where each respondent's elective choices were treated as a transaction. The elective courses selected by each student formed the items in the transaction.

We then applied the Apriori algorithm, a widely-used association rule mining technique, to the transactional data. This algorithm allowed us to extract association rules that show the likelihood of certain electives being chosen together based on academic tracks.

Methodology

Throughout the experiment, we strictly adhered to ethical guidelines to protect the privacy and rights of the participants. The data was anonymized and aggregated, ensuring that individual identities remained confidential.

Any personal data collected during the survey was handled with utmost confidentiality and solely for research purposes.

Experiments results

For every track we can see what elective they will likely chose which will help us understand do they chose their elective based on it's content which is suitable for their track or just choose a random one.

Experiments results

Let's take a look at Pic. 2.

From here we can see that students from that student mostly choose Fronted and programming languages, but don't choose something regarding to math, devops or AI.

	antecedents		consequents	confidence
70	(Track_SD)		(Tech22_Front-end Web Development)	0.316456
86	(Track_SD)		(Tech23_Programming in Python)	0.291139
72	(Track_SD)		(Tech22_Functional Programming in Haskell)	0.227848
66	(Track_SD)	(Tech22_Computer Graphics in Game Development)		0.202532
76	(Track_SD)	(Tech23_Architecture of high-load systems)		0.139241
68	(Track_SD)	(Tech22_Cross-platform Mobile Development with...		0.126582
88	(Track_SD)		(Tech23_System Programming)	0.113924
82	(Track_SD)		(Tech23_Front-end Web Development)	0.101266
90	(Track_SD)		(Tech23_iOS Development with Swift)	0.088608
84	(Track_SD)		(Tech23_Introduction to DevOps)	0.075949
78	(Track_SD)	(Tech23_Challenges of Object Oriented Programm...		0.063291

Pic. 2 Recommendation for SD on Tech courses

Experiments results

On the other hand, we can look at the Pic. 3. We can see that CS students choose something related to Linux and security.

Which means students choose their tech electives depending on their track

	antecedents		consequents	confidence
21	(Track_CS)	(Tech22_Functional Programming in Haskell)		0.346154
33	(Track_CS)		(Tech23_Introduction to DevOps)	0.269231
29	(Track_CS)	(Tech23_Blockchain: Distributed Ledger Develop...		0.192308
19	(Track_CS)	(Tech22_Cross-platform Mobile Development with...		0.192308
35	(Track_CS)		(Tech23_System Programming)	0.192308
25	(Track_CS)	(Tech22_Lambda-Calculus, Algebra, Machinery an...		0.115385
31	(Track_CS)		(Tech23_Front-end Web Development)	0.115385
27	(Track_CS)		(Tech23_Architecture of high-load systems)	0.115385
23	(Track_CS)	(Tech22_Introduction to Robotics Operating Sys...		0.076923

Pic. 3 Recommendation for CS on Tech courses

Experiments results

In humanitarian electives we can see some connection, but it is hard to say was it depending on track or just to find an easy elective.

	antecedents		consequents	confidence
90	(Track_SD)	(Hum22_Introduction to Career Development for ...		0.202532
96	(Track_SD)	(Hum23_Basics of Product Management)		0.189873
84	(Track_SD)	(Hum22_Career and leadership)		0.164557
93	(Track_SD)	(Hum22_Psychology of IT-specialist)		0.151899
98	(Track_SD)	(Hum23_Design thinking)		0.126582
82	(Track_SD)	(Hum22_Applied Economics: Introduction to IT E...		0.126582
102	(Track_SD)	(Hum23_Introduction to Public Speaking for IT-...		0.126582
108	(Track_SD)	(Hum23_Startup CEO Toolkit)		0.113924
106	(Track_SD)	(Hum23_Social Entrepreneurship)		0.113924
104	(Track_SD)	(Hum23_Personal Efficiency Skills of IT-specia...		0.101266
86	(Track_SD)	(Hum22_Design Fiction)		0.101266
88	(Track_SD)	(Hum22_Developing Entrepreneurial Skills for a...		0.088608
100	(Track_SD)	(Hum23_Introduction to Career Development for ...		0.075949
94	(Track_SD)	(Hum22_Tech Startup Design)		0.075949

Pic. 4 Recommendation for SD on Hum courses

	antecedents		consequents	confidence
71	(Track_DS)	(Hum23_Psychology of IT-specialist)		0.265306
48	(Track_DS)	(Hum22_Applied Economics: Introduction to IT E...		0.183673
50	(Track_DS)	(Hum22_Design Fiction)		0.183673
56	(Track_DS)	(Hum22_Personal Efficiency Skills of IT-specia...		0.183673
54	(Track_DS)	(Hum22_Introduction to Career Development for ...		0.163265
65	(Track_DS)	(Hum23_Introduction to Career Development for ...		0.163265
53	(Track_DS)	(Hum22_Developing Entrepreneurial Skills for a...		0.102041
67	(Track_DS)	(Hum23_Introduction to Public Speaking for IT-...		0.102041
63	(Track_DS)	(Hum23_How to build an IT team)		0.102041
61	(Track_DS)	(Hum23_Design thinking)		0.102041
59	(Track_DS)	(Hum22_Psychology of IT-specialist)		0.102041
72	(Track_DS)	(Hum23_Social Entrepreneurship)		0.081633
69	(Track_DS)	(Hum23_Personal Efficiency Skills of IT-specia...		0.081633

Pic. 4 Recommendation for DS on Hum courses

Result of the internship

At the result I wanted to give a suggestion for elective courses distribution, but from the experiment I saw that everything is mostly fine, and we can consider a plus point for those who are making this distributions.

Also I will leave my code, which is scalable and can be reused if there will be better data.