

Určovanie **zvedavosti** používateľov z ich správania v doméne online obchodu

Jakub Šimko
jakub.simko@stuba.sk

Sandra Kostova
kostova.sandra@gmail.com

Mária Bieliková
maria.bielikova@stuba.sk



Jeszika Le Vye: Curiosity



PeWe@FIIT
personalized web group



SLOVAK UNIVERSITY OF
TECHNOLOGY IN BRATISLAVA
FACULTY OF INFORMATICS
AND INFORMATION TECHNOLOGIES

Miera zvedavosti ovplyvňuje správanie človeka



**Spotrebiteľské
správanie...**



...aj na webe



**Ponúka
vysvetlenia**

Zvedavosť je merateľná dotazníkom

(Curiosity and exploration inventory)

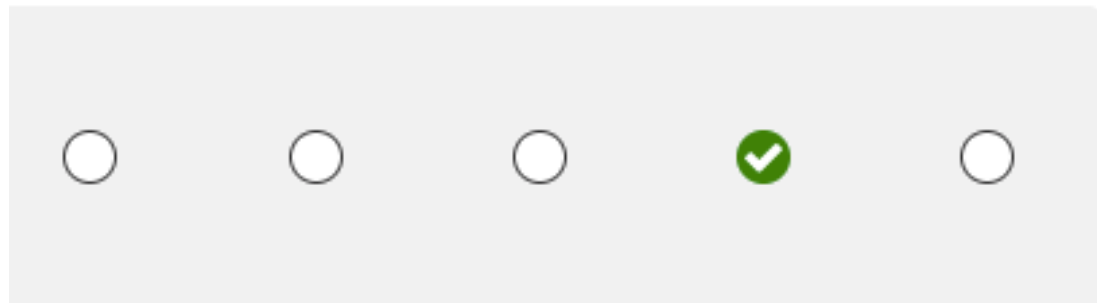
Seba-hodnotiaci

10 výrokov

Príklad výroku:

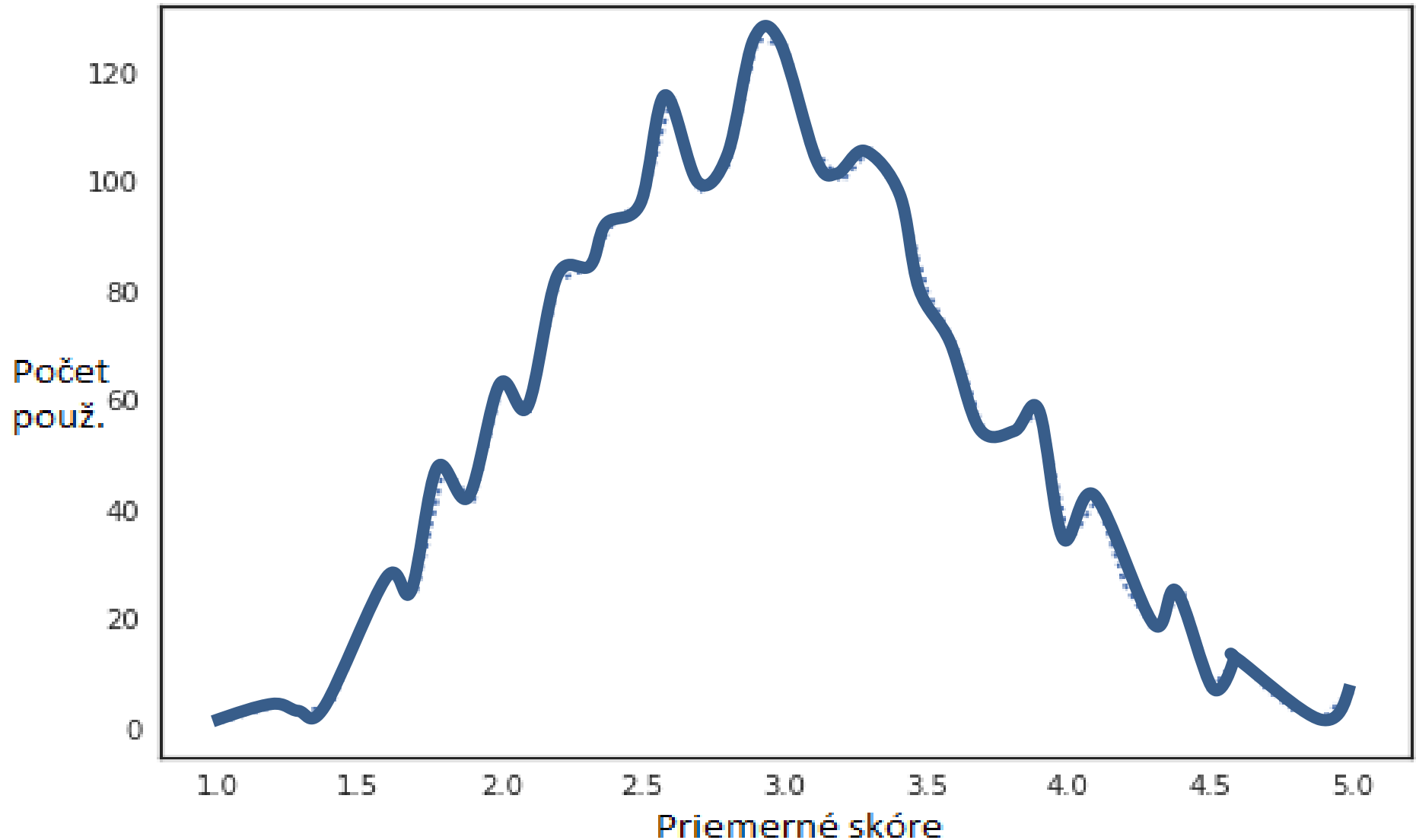
„Všade kde chodím, obzerám sa po nových veciach a zážitkoch.“

Strongly Disagree Disagree Neutral Agree Strongly Agree

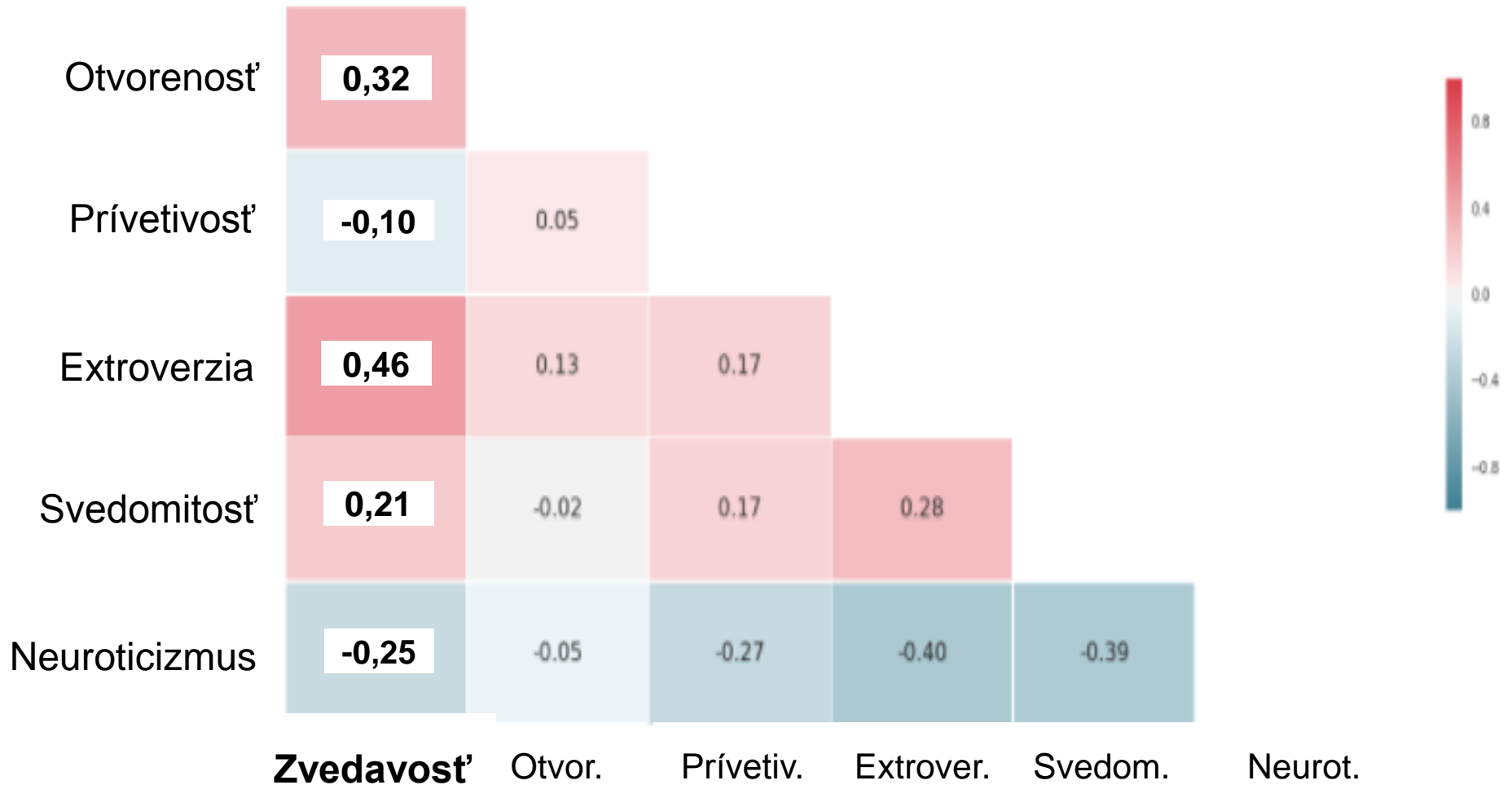


A horizontal Likert scale with five radio buttons. The fourth button, labeled 'Agree', is selected and highlighted with a green checkmark.

V našej vzorke 2130 používateľov online obchodu je miera zvedavosti **normálna**



Zvedavosť má vzťah s charakteristikami **big five**



Zaujímalo nás, či vieme mieru zvedavosti odvodit' zo správania na webe



```
2008-10-09 00:37:39.152 CHSVSR07 21384 12764 SW_SQLDataOutputHandler::open - - - 2 41494 "Failed to establish a connection to the data source: (Server: CHSSQL01, Database: ...)"
2008-10-09 00:37:39.158 - 18568 18792 SW_MessageHandler::close - - - 5 65535 "Initialization complete."
2008-10-09 00:37:39.617 - 18900 20176 SW_MessageHandler::open - - - 5 65535 "Initialization complete."
2008-10-09 00:37:39.668 - 21572 17944 SW_MessageHandler::open - - - 5 65535 "Initialization complete."
2008-10-09 00:37:40.308 CHSVSR07 18900 20176 SW_SQLDataOutputHandler::open - - - 1 44910 "Failed to establish a connection to the data source: (Server: CHSSQL01, Database: ...)"
2008-10-09 00:37:40.308 CHSVSR07 18900 20176 SW_SQLOutputHandler::open - - - 1 44910 "Failed to connect to data source."
2008-10-09 00:37:40.995 CHSVSR07 21572 17944 SW_SQLDataConnection::open - - - 2 41494 "Failed to establish a connection to the data source: (Server: CHSSQL01, Database: ...)"
2008-10-09 00:37:41.012 CHSVSR07 21572 17944 SW_SQLOutputHandler::open - - - 1 44910 "Failed to connect to data source."
2008-10-09 00:37:41.215 CHSVSR07 18568 18792 SW_SQLDataConnection::open - - - 2 41494 "Failed to establish a connection to the data source: (Server: CHSSQL01, Database: ...)"
2008-10-09 00:37:41.215 CHSVSR07 18568 18792 SW_SQLOutputHandler::open - - - 1 44910 "Failed to connect to data source."
2008-10-09 00:37:41.980 CHSVSR07 21384 12764 SW_SystemDispatcher::Init - - - 0 44952 "Successfully started Microsoft System Center Application Virtualization Management"
2008-10-09 00:37:41.980 CHSVSR07 21384 12764 SW_SystemDispatcher::Final - - - 0 44952 "Successfully shut down Microsoft System Center Application Virtualization Management"
2008-10-09 00:36:32.292 CHSVSR07 21384 22340 SW_MessageHandler::close - - - 5 65535 "Shutdown complete."
2008-10-09 00:36:32.573 CHSVSR07 18568 18792 SW_MessageHandler::close - - - 5 65535 "Shutdown complete."
2008-10-09 00:36:32.573 CHSVSR07 18900 20176 SW_MessageHandler::close - - - 5 65535 "Shutdown complete."
2008-10-09 00:36:32.573 CHSVSR07 21572 17944 SW_MessageHandler::close - - - 5 65535 "Shutdown complete."
2008-10-09 00:58:37.375 - 1140 1104 SW_MessageHandler::open - - - 5 65535 "Initialization complete."
2008-10-09 00:58:41.245 CHSVSR07 1140 1104 SW_SQLDataConnection::open - - - 2 41494 "Failed to establish a connection to the data source: (Server: CHSSQL01, Database: ...)"
2008-10-09 00:58:41.245 CHSVSR07 1140 1104 SW_SQLOutputHandler::open - - - 1 44910 "Failed to connect to data source."
2008-10-09 00:58:41.765 - 1492 1496 SW_MessageHandler::open - - - 5 65535 "Initialization complete."
2008-10-09 00:58:41.765 - 1504 1508 SW_MessageHandler::open - - - 5 65535 "Initialization complete."
2008-10-09 00:58:41.234 CHSVSR07 1504 1508 SW_SQLDataConnection::open - - - 2 41494 "Failed to establish a connection to the data source: (Server: CHSSQL01, Database: ...)"
2008-10-09 00:58:41.234 CHSVSR07 1504 1508 SW_SQLOutputHandler::open - - - 1 44910 "Failed to connect to data source."
2008-10-09 00:58:41.234 CHSVSR07 1504 1508 SW_SQLDataConnection::open - - - 2 41494 "Failed to establish a connection to the data source: (Server: CHSSQL01, Database: ...)"
2008-10-09 00:58:41.234 CHSVSR07 1504 1508 SW_SQLOutputHandler::open - - - 1 44910 "Failed to connect to data source."
2008-10-09 00:58:41.319 CHSVSR07 1492 1496 SW_SQLDataConnection::open - - - 2 41494 "Failed to establish a connection to the data source: (Server: CHSSQL01, Database: ...)"
2008-10-09 00:58:41.319 CHSVSR07 1492 1496 SW_SQLOutputHandler::open - - - 1 44910 "Failed to connect to data source."
2008-10-09 00:58:46.228 CHSVSR07 1140 1104 SW_SystemDispatcher::Init - - - 0 44952 "Successfully started Microsoft System Center Application Virtualization Management"
2008-10-09 00:58:46.228 CHSVSR07 1140 1104 SW_SystemDispatcher::Final - - - 0 44952 "Successfully shut down Microsoft System Center Application Virtualization Management"
2008-10-09 01:15:25.370 CHSVSR07 1140 1104 SW_MessageHandler::close - - - 5 65535 "Shutdown complete."
2008-10-09 01:15:25.940 CHSVSR07 1492 1496 SW_MessageHandler::close - - - 5 65535 "Shutdown complete."
2008-10-09 01:15:26.632 CHSVSR07 1512 1516 SW_MessageHandler::close - - - 5 65535 "Shutdown complete."
2008-10-09 01:15:26.694 CHSVSR07 1504 1508 SW_MessageHandler::close - - - 5 65535 "Shutdown complete."
2008-10-09 01:15:27.395 - 5320 5324 SW_MessageHandler::open - - - 5 65535 "Initialization complete."
2008-10-09 01:15:28.673 CHSVSR07 5192 5384 SW_SQLDataConnection::open - - - 2 41494 "Failed to establish a connection to the data source: (Server: CHSSQL01, Database: ...)"
2008-10-09 01:15:28.673 CHSVSR07 5192 5384 SW_SQLOutputHandler::open - - - 1 44910 "Failed to connect to data source."
2008-10-09 01:15:29.016 - 6400 7732 SW_MessageHandler::open - - - 5 65535 "Initialization complete."
2008-10-09 01:15:29.079 - 5732 6148 SW_MessageHandler::open - - - 5 65535 "Initialization complete."
2008-10-09 01:15:29.954 CHSVSR07 6400 7732 SW_SQLDataConnection::open - - - 2 41494 "Failed to establish a connection to the data source: (Server: CHSSQL01, Database: ...)"
2008-10-09 01:15:29.904 CHSVSR07 6400 7732 SW_SQLOutputHandler::open - - - 1 44910 "Failed to connect to data source."
2008-10-09 01:15:30.278 CHSVSR07 7432 6396 SW_SQLDataConnection::open - - - 2 41494 "Failed to establish a connection to the data source: (Server: CHSSQL01, Database: ...)"
2008-10-09 01:15:30.284 CHSVSR07 7432 6396 SW_SQLOutputHandler::open - - - 1 44910 "Failed to connect to data source."
2008-10-09 01:15:30.356 CHSVSR07 5732 6148 SW_SQLDataConnection::open - - - 2 41494 "Failed to establish a connection to the data source: (Server: CHSSQL01, Database: ...)"
2008-10-09 01:15:30.356 CHSVSR07 5732 6148 SW_SQLOutputHandler::open - - - 1 44910 "Failed to connect to data source."
2008-10-09 01:15:31.185 CHSVSR07 5872 5884 SW_SystemDispatcher::Init - - - 0 44952 "Successfully started Microsoft System Center Application Virtualization Management"
2008-10-09 01:15:31.185 CHSVSR07 5872 5884 SW_SystemDispatcher::Final - - - 0 44952 "Successfully shut down Microsoft System Center Application Virtualization Management"
2008-10-09 01:15:36.785 CHSVSR07 7432 8120 SW_LicenseConduitLogger::LogMessage 1956138080 "Default Provider" Justin Zarb TECHDES.001/TECHDES.001.sft 0 40976 "License S"
2008-10-09 01:15:36.911 CHSVSR07 7432 8120 SW_LicenseConduitLogger::LogMessage 1956138080 "Default Provider" Justin Zarb TECHDES.001/TECHDES.001.sft 0 40960 "Session Setup"
2008-10-09 01:16:03.926 CHSVSR07 7432 8182 SW_LicenseConduitLogger::LogMessage 1956138080 "Default Provider" Justin Zarb TECHDES.001/TECHDES.001.sft 0 40977 "License S"
2008-10-09 01:16:03.926 CHSVSR07 7432 8182 SW_LicenseConduitLogger::LogMessage 1956138080 "Default Provider" Justin Zarb TECHDES.001/TECHDES.001.sft 0 40961 "Session Teard"
2008-10-09 01:20:32.746 CHSVSR07 6400 1844 SW_LicenseConduitLogger::LogMessage 1897464568 "Default Provider" Justin Zarb SCVMCON.001/SCVMCON.001.sft 0 40960 "Session Setup"
2008-10-09 01:20:32.746 CHSVSR07 6400 1844 SW_LicenseConduitLogger::LogMessage 1897464568 "Default Provider" Justin Zarb SCVMCON.001/SCVMCON.001.sft 0 40960 "Session Setup"
2008-10-09 06:11:40.612 CHSVSR07 6400 8016 SW_LicenseConduitLogger::LogMessage 1897464568 "Default Provider" Justin Zarb SCVMCON.001/SCVMCON.001.Package 0 40978 "License Abnorm"
2008-10-09 06:11:40.663 CHSVSR07 6400 8340 SW_SQLOutputHandler::HandleMessage - - - 1 44811 "Create record failed with error [1]."
2008-10-09 06:35:40.367 CHSVSR07 5732 3960 SW_DataConnectionPool::Reconnect - "Default Provider" Justin Zarb SCVMCON.001/SCVMCON.001.sft 2 41543 "Connection to data"
2008-10-09 06:35:40.367 CHSVSR07 5732 3960 SW_ServerAuthenticationTasks::Authorize - "Default Provider" Justin Zarb SCVMCON.001/SCVMCON.001.sft 2 41472 "Module execut"
2008-10-09 06:35:40.367 CHSVSR07 5732 3960 SW_SQLOutputHandler::HandleMessage - - - 1 44811 "Create record failed with error [1]."
2008-10-09 06:48:44.656 CHSVSR07 7432 6572 SW_DataConnectionPool::Reconnect - "Default Provider" ecoe - 2 41543 "Connection to data store on host CHSSQL01/SSQSTEPS fail"
2008-10-09 06:48:44.656 CHSVSR07 7432 6572 SW_ServerAuthenticationTasks::Authorize - "Default Provider" ecoe - 2 41472 "Module execution failed."
2008-10-09 06:48:44.656 CHSVSR07 6400 1872 SW_DataConnectionPool::Reconnect - "Default Provider" ecoe - 2 41543 "Connection to data store on host CHSSQL01/SSQSTEPS fail"
2008-10-09 06:48:44.656 CHSVSR07 7432 7864 SW_SQLOutputHandler::HandleMessage - - - 1 44811 "Create record failed with error [1]."
2008-10-09 06:48:44.702 CHSVSR07 6400 5872 SW_ServerAuthenticationTasks::Authorize - "Default Provider" ecoe - 2 41472 "Module execution failed."
2008-10-09 08:13:07.803 CHSVSR07 5732 3960 SW_LicenseConduitLogger::LogMessage 2111456567 "Default Provider" Swishman S2MSNET.005/S2MSNET.005.7.sft 0 40976 "License Sess"
2008-10-09 08:13:07.803 CHSVSR07 5732 3960 SW_LicenseConduitLogger::LogMessage 2111456567 "Default Provider" Swishman S2MSNET.005/S2MSNET.005.7.sft 0 40960 "Session Setup"
2008-10-09 08:13:07.818 CHSVSR07 5732 6508 SW_LicenseConduitLogger::LogMessage 2111456567 "Default Provider" Swishman S2MSNET.005/S2MSNET.005.7.sft 0 40960 "Session Setup"
2008-10-09 08:13:07.818 CHSVSR07 5732 3960 SW_LicenseConduitLogger::LogMessage 2056184000 "Default Provider" Swishman S2MSNET.005/S2MSNET.005.7.sft 0 40976 "License Sess"
2008-10-09 08:13:07.818 CHSVSR07 5732 3960 SW_LicenseConduitLogger::LogMessage 2056184000 "Default Provider" Swishman S2MSNET.005/S2MSNET.005.7.sft 0 40960 "Session Setup"
2008-10-09 08:44:58.679 CHSVSR07 7432 6492 SW_LicenseConduitLogger::LogMessage 100801806 "Default Provider" DLANEY ADDR811.001/ADDR811.001.sft 0 40976 "License Sessio"
2008-10-09 08:44:58.679 CHSVSR07 7432 6492 SW_LicenseConduitLogger::LogMessage 100801806 "Default Provider" DLANEY ADDR811.001/ADDR811.001.sft 0 40960 "Session Setup"
2008-10-09 08:45:03.412 CHSVSR07 7432 5528 SW_LicenseConduitLogger::LogMessage 100801806 "Default Provider" DLANEY ADDR811.001/ADDR811.001.sft 0 40977 "License Sessio"
```

Dataset s vyše 300 000 udalost'ami (slovenský e-shop)

Analyzovali sme 45 potenciálnych črt

Demografia (vek, pohlavie, príjmová skupina, ...)

Celková aktivita používateľov (počty udalostí podľa typov, sedení, časy, časti dňa, ...)

Spôsob začiatku sedenia (titulka, soc. sieť, vyhľadávač, iná stránka, ...)

Informácie o sedeniach (trvanie, počet, rozostup medzi...)

N-gramy (produkt-produkt, kategória-produkt, kategória-kategória, ...)

Práca s kategóriami (priemerný počet počas sedenia, miera prezerania podkategórií, ...)

Práca so zoznamami produktov (počet návštev druhej strany výsledkov vyhľadávania, ...)

Klasifikovali sme **náhodným lesom do 3 tried** (veľmi, priemerne a málo zvedaví)

	Acc.	Prec.	Rec.
Baseline	0,44		
Všetky črty (45)	0,66	0,56	64,5
Feature importance (7)	0,65		
PCA (11)	0,65		
Príchod zo soc. siete + uskutočnenie nákupu	0,70	0,68	0,49

Zvedavosť sa odráža na správaní na webe a aj spätne ju možno (do nejakej miery) odhadnúť

Dataset správania v e-shope:

2130 používateľov

300 000 udalostí

Skupiny črt:

Demografia

Celková aktivita používateľov

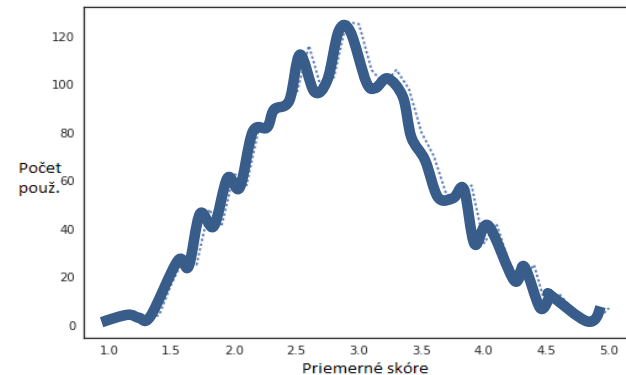
Spôsob začiatku sedenia

Informácie o sedeniach

N-gramy Práca s kategóriami

Práca so zoznamami produktov

Referenčná informácia o zvedavosti



Presnosť natrénovaných modelov

	Acc.
Všetky črty (45)	0,66
Príchod zo soc. siete + uskutočnenie nákupu	0,70