

Tato prezentace je součástí wiki-prezentace [Metoda GUHA, LISp-Miner a typové úlohy](#)

Je dostupná z [této adresy](#)

Verze 13. 10. 2019

Typ úlohy: Využití poměru konfidencí u SD4ft-pravidel

Data: [Accidents](#)

Problém: *Porovnání mužů a žen co se týče relativní četnosti nehod s různou závažností pro různé kombinace charakteristiky řidiče a okolností nehody.*

Jan Rauch

Katedra informačního a znalostního inženýrství

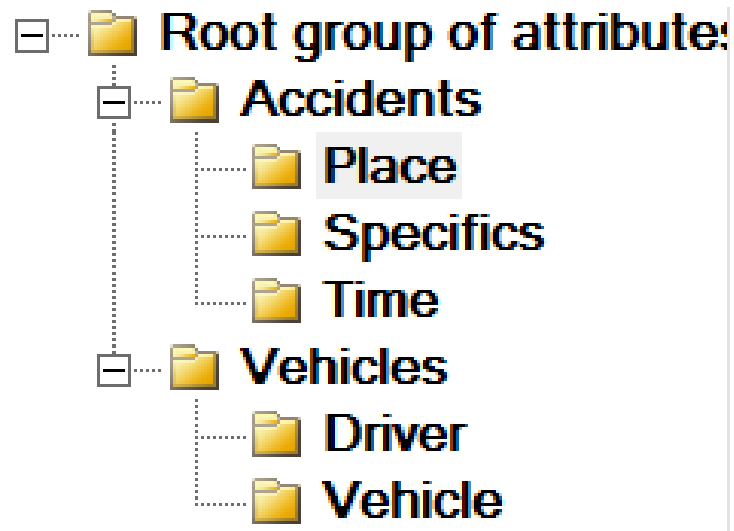
Vysoká škola ekonomická v Praze

Procedura SD4ft-Miner - využití poměru konfidencí

- Data Accidents
- Vysoký poměr konfidencí
- Poměr konfidencí přibližně jedna

Data Accidents - přehled

- Jedná se o data vzniklá v rámci studentské seminární práce v předmětu *4IZ460 - pokročilé přístupy k DZD* v zimním semestru 2018/2019 na VŠE Praha.
- Výchozí data se skládají ze tří souborů dostupných na <https://www.kaggle.com/silicon99/dft-accident-data>
- Byly zpracovány pouze nehody jednoho vozidla vzhledem k jasné odpovědnosti řidiče.
- Vzniklá matice dat má 538 989 řádků
- Vytvořené atributy rozděleny do skupin:



Data Accidents - atributy (1)

Groups of attributes tree	Attribute	Used	DBCColumn	Categories	XCal	Sample categories
<ul style="list-style-type: none"> Root group of attribute: <ul style="list-style-type: none"> Accidents <ul style="list-style-type: none"> Place 	Area_type	+	Urban_or_Rural_An	3		Urban, Rural, Unallocated

Groups of attributes tree	Attribute	Used	DBCColumn	Categories	XCal	Sample categories
<ul style="list-style-type: none"> Root group of attribute: <ul style="list-style-type: none"> Accidents <ul style="list-style-type: none"> Place <ul style="list-style-type: none"> Specifics Time Vehicles <ul style="list-style-type: none"> Driver Vehicle 	1st_Road_Class	+	x_1st_Road_Class	6		Motorway, A(M), A, B, C, Unclassified
	Accident_Severity	+	Accident_Severity	3		Fatal, Serious, Slight
	Carriageway_Hazards	+	Carriageway_Hazar	6	x	None, Vehicle load on road, Other object on road, Previous accident, Pedestrian
	Carriageway_Hazards 5 000+	+	Carriageway_Hazar	3	x	None, Other object on road, Any animal in carriageway
	Did_Police_Officer_Attend_Scene_of_Acc	+	Did_Police_Officer_	3	x	Yes, No, Self rep only
	Did_Police_Officer_Attend_Scene_of_Acc +	+	Did_Police_Officer_	2	x	Yes, No
	Junction_Detail	+	Junction_Detail	9	x	Not at junction or within 20 metres, Roundabout, Mini-roundabout, T or staggered
	Light_Conditions	+	Light_Conditions	5		Daylight, Darkness - lights lit, Darkness - lights unlit, Darkness - no lighting, Dark
	Number_of_Casualties		Number_of_Casualt	33		1, 2, 3, 4, 5, 6, 7, 8, 9, 10...
	Number_of_Casualties_ef5		Number_of_Casualt	5		<1;2>, <2;3>, <3;4>, <4;5>, <5;68>
	Number_of_Casualties_ef5 5000 +	+	Number_of_Casualt	3		<1;2>, <2;3>, <3;4>
	Road_Surface_Conditions	+	Road_Surface_Con	5	x	Dry, Wet or damp, Snow, Frost or ice, Flood over 3cm. deep
	Road_Surface_Conditions 5000 +	+	Road_Surface_Con	3	x	Dry, Wet or damp, Frost or ice
	Speed_limit	+	Speed_limit	8		10, 15, 20, 30, 40, 50, 60, 70
	Speed_limit_ed3	+	Speed_limit	3		<10;29>, <30;49>, <50;70>
	Weather_Conditions	+	Weather_Conditions	9	x	Fine no high winds, Raining no high winds, Snowing no high winds, Fine + high v
	Weather_Conditions 5 000 +	+	Weather_Conditions	8	x	Fine no high winds, Raining no high winds, Snowing, Fine + high winds, Raining

Groups of attributes tree	Attribute	Used	DBCColumn	Categories	XCal	Sample categories
<ul style="list-style-type: none"> Root group of attribute: <ul style="list-style-type: none"> Accidents <ul style="list-style-type: none"> Place Specifics Time Vehicles <ul style="list-style-type: none"> Driver Vehicle 	Day_of_Week	+	Day_of_Week	7		Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday
	Hour	+	Time_Hour	24	x	0, 1, 2, 3, 4, 5, 6, 7, 8, 9...
	Month	+	Date_Month	12		January, February, March, April, May, June, July, August, September, October.
	Quarter	+	Date_Quarter	4		Q1, Q2, Q3, Q4
	Rush_hours	+	Time_Hour	2	x	Other, 8-9/16-18
	Weekend	+	Day_of_Week	2		Weekend, Weekdays
	Year	+	Date_Year	11		2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014...

Data Accidents - atributy (2)

Groups of attributes tree	Attribute	Used	DBCColumn	Categories	XCal	Sample categories
[-] Root group of attribute:	Age_Band_of_Driver		Age_Band_of_Drive	11	x	0 - 5, 6 - 10, 11 - 15, 16 - 20, 21 - 25, 26 - 35, 36 - 45, 46 - 55, 56 - 65, 66 - 75...
[-] Accidents	Age_Groups_of_Driver	+	Age_of_Driver	9	x	Under 18, 18 to 25, 26 to 35, 36 to 45, 46 to 55, 56 to 65, 66 to 75, 76 to 85, Ab
[-] Place	Age_of_Driver		Age_of_Driver	98	x	1, 2, 3, 4, 5, 6, 7, 8, 9, 10...
[-] Specifics	Age_of_Driver_ef20		Age_of_Driver	20		ef_1, ef_2, ef_3, ef_4, ef_5, ef_6, ef_7, ef_8, ef_9, ef_10...
[-] Time	Age_of_Driver_ef5	+	Age_of_Driver	5	x	<1;22>, <22;30>, <30;40>, <40;52>, <52;98>
[-] Vehicles	Age_of_Driver_mod		Age_of_Driver	98		1, 2, 3, 4, 5, 6, 7, 8, 9, 10...
[-] Driver	Driver_Home_Area_Type	+	Driver_Home_Area_	3	x	Urban area, Small town, Rural
[-] Vehicle	Driver_IMD_Decile	+	Driver_IMD_Decile	10	x	Most deprived 10%, More deprived 10-20%, More deprived 20-30%, More depriv
	Journey_Purpose_of_Driver	+	Journey_Purpose_o	5	x	Journey as part of work, Commuting to/from work, Taking pupil to/from school, P
	Sex_of_Driver	+	Sex_of_Driver	2	x	Male, Female

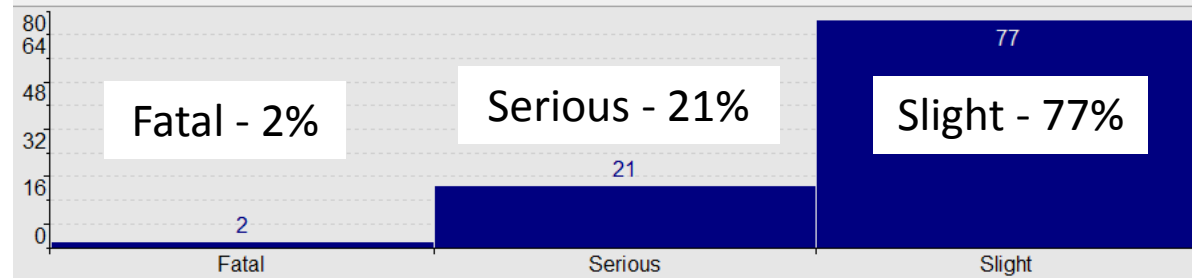
Groups of attributes tree	Attribute	Used	DBCColumn	Categories	XCal	Sample categories
[-] Root group of attribute:	Age_of_Vehicle	+	Age_of_Vehicle	82	x	1, 2, 3, 4, 5, 6, 7, 8, 9, 10...
[-] Accidents	Age_of_Vehicle_ed5	+	Age_of_Vehicle	21	x	<0;4>, <5;9>, <10;14>, <15;19>, <20;24>, <25;29>, <30;34>, <35;39>, <40;44>,
[-] Place	Age_of_Vehicle_ef	+	Age_of_Vehicle	9	x	<2;3>, <3;5>, <5;6>, <6;7>, <7;8>, <8;10>, <10;11>, <11;13>, <13;105>
[-] Specifics	Age_of_Vehicle_ef20	+	Age_of_Vehicle	20		ef_1, ef_2, ef_3, ef_4, ef_5, ef_6, ef_7, ef_8, ef_9, ef_10...
[-] Time	Age_of_Vehicle_mod	+	Age_of_Vehicle	82		1, 2, 3, 4, 5, 6, 7, 8, 9, 10...
[-] Vehicles	Engine_Capacity__CC_	+	Engine_Capacity__(5	x	<4;1199>, <1199;1398>, <1398;1794>, <1794;1998>, <1998;99999>
[-] Driver	Propulsion_Code	+	Propulsion_Code	11	x	Petrol, Heavy oil, Electric, Steam, Gas, Petrol/Gas (LPG), Gas/Bi-fuel, Hybrid ele
[-] Vehicle	Propulsion_Code 5000+	+	Propulsion_Code	2	x	Petrol, Heavy oil
	Vehicle_Type	+	Vehicle_Type	20	x	Pedal cycle, Motorcycle 50cc and under, Motorcycle 125cc and under, Motorcyc
	Vehicle_Type 5000 +	+	Vehicle_Type	10	x	Pedal cycle, Motorcycle 50cc and under, Motorcycle 125cc and under, Motorcyc
	Vehicle_Type_Groups	+	Vehicle_Type	8	x	Pedal cycle, Motorcycle, Car, Bus, Other, Agricultural vehicle, Tram, Van
	Vehicle_Type_Groups 5 000 +	+	Vehicle_Type	5	x	Pedal cycle, Motorcycle, Car, Bus, Van
	Was_Vehicle_Left_Hand_Drive	+	Was_Vehicle_Left_	2	x	No, Yes

Procedura SD4ft-Miner - využití poměru konfidencí

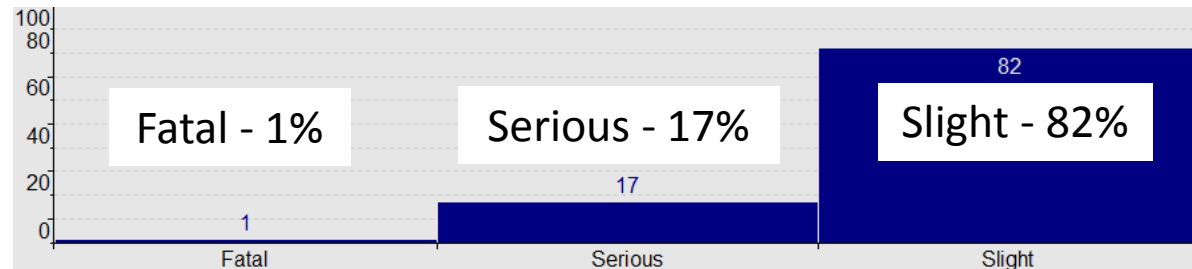
- Data Accidents
- Vysoký poměr konfidencí
- Poměr konfidencí přibližně jedna

Vysoký poměr konfidencí - inspirace

Procentuální podíl hodnot atributu Accident_Severity pro řidiče - muže:



Procentuální podíl hodnot atributu Accident_Severity pro řidiče - ženy:



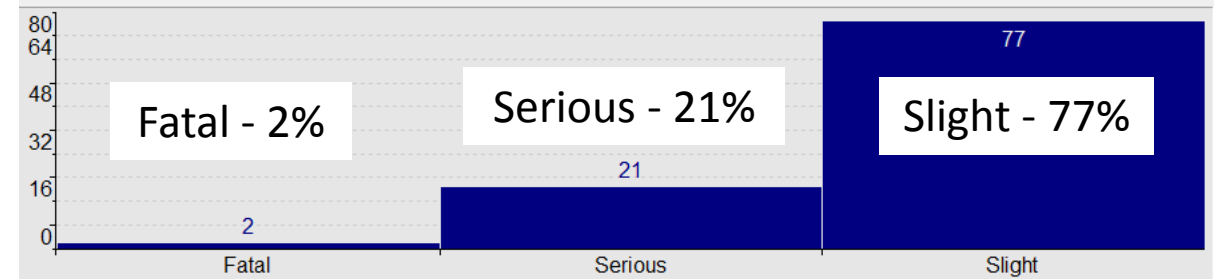
Pro muže jsou fatální nehody 2 x častější než pro ženy

Otázka:

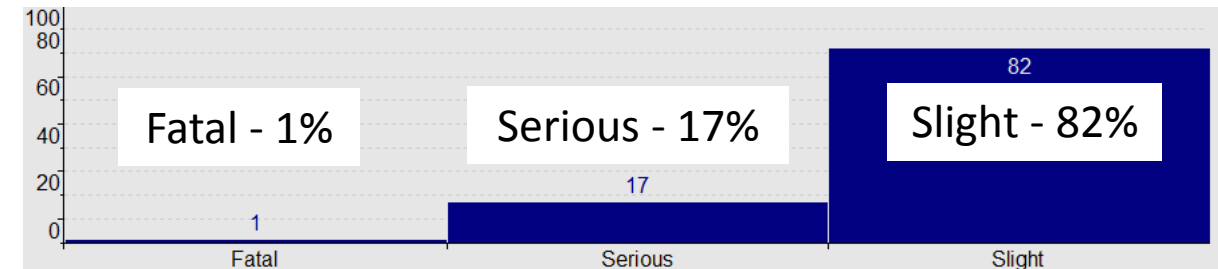
Je pro některé řidiče za nějakých okolností relativní četnost poruch s nějakou závažností víc než dvakrát větší / menší pro muže než pro ženy?

Analytická otázka

Konfidence: True \approx Severity(Fatal)/Sex(Male) = 0.02



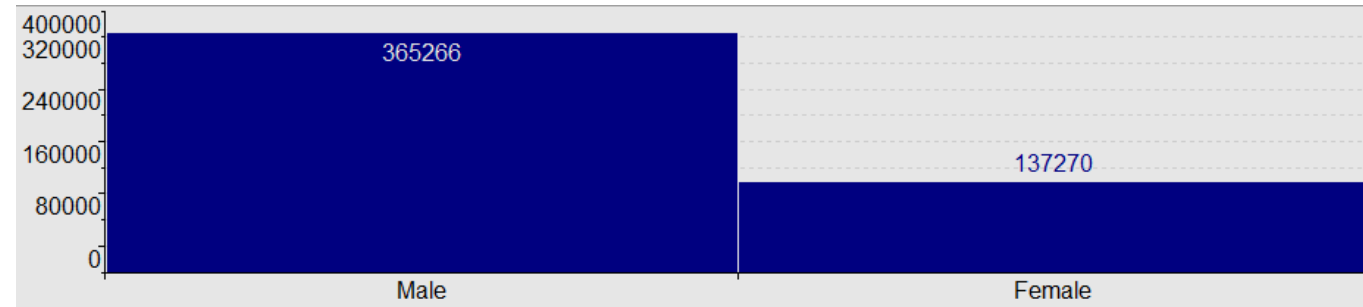
Konfidence: True \approx Severity(Fatal)/ Sex(Female) = 0.01



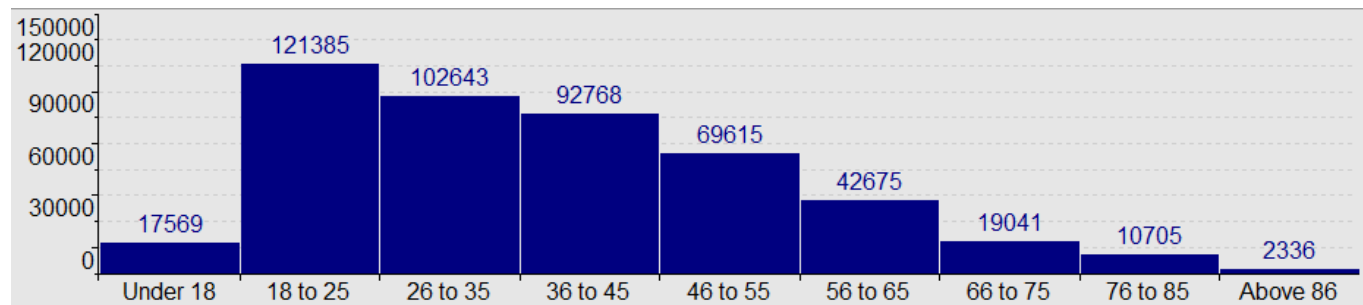
$$\frac{\text{Konfidence Driver(?) } \wedge \text{ Circumstancies(?) } \approx \text{ Severity(?) / Sex(?)}}{\text{Konfidence Driver(?) } \wedge \text{ Circumstancies(?) } \approx \text{ Severity(?) / Sex(?)}} > 2$$

Data o řidiči

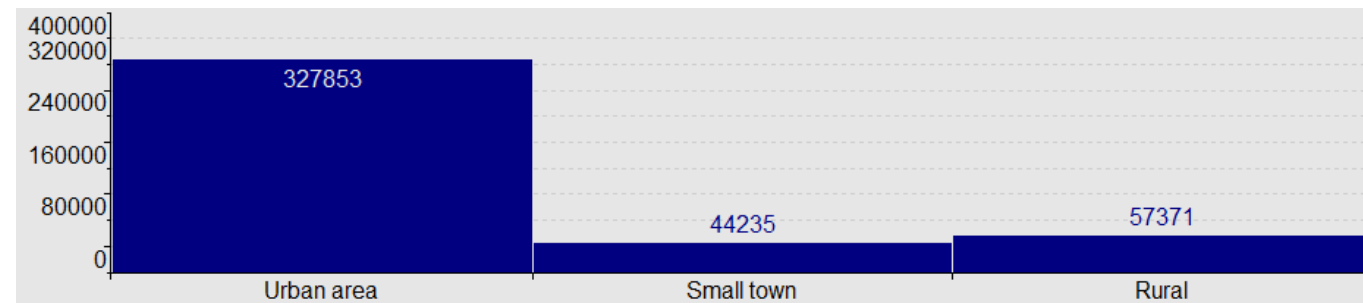
Sex_of_Driver



Age_Groups_of_Driver

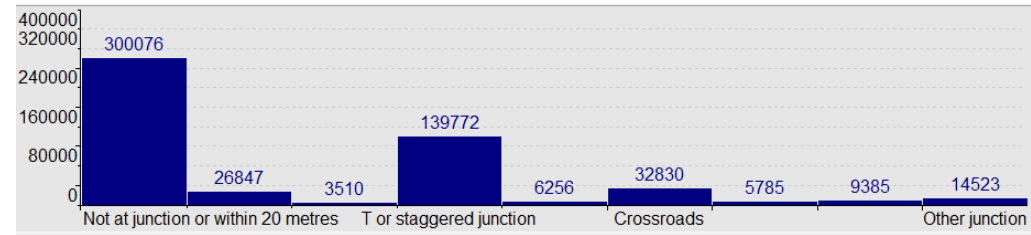


Driver_Home_Area_Type

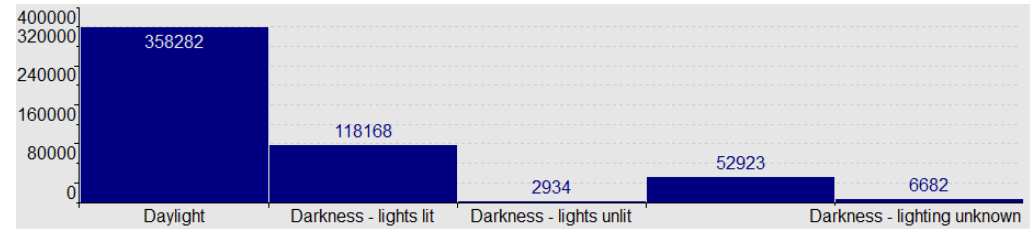


Data o podmínkách nehody

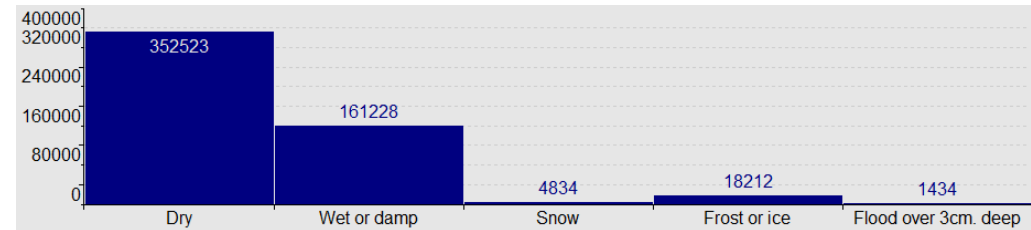
Junction_Detail



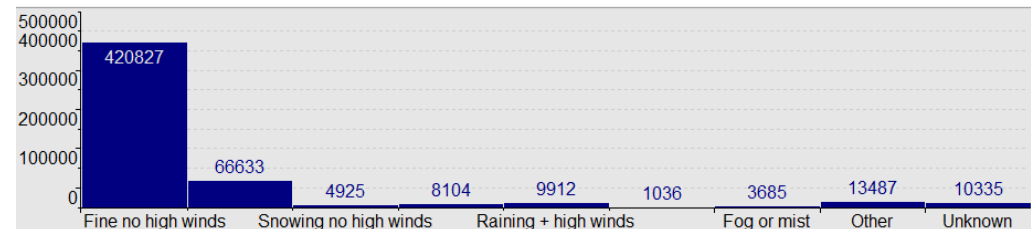
Light_Conditions



Road_Surface_Conditions

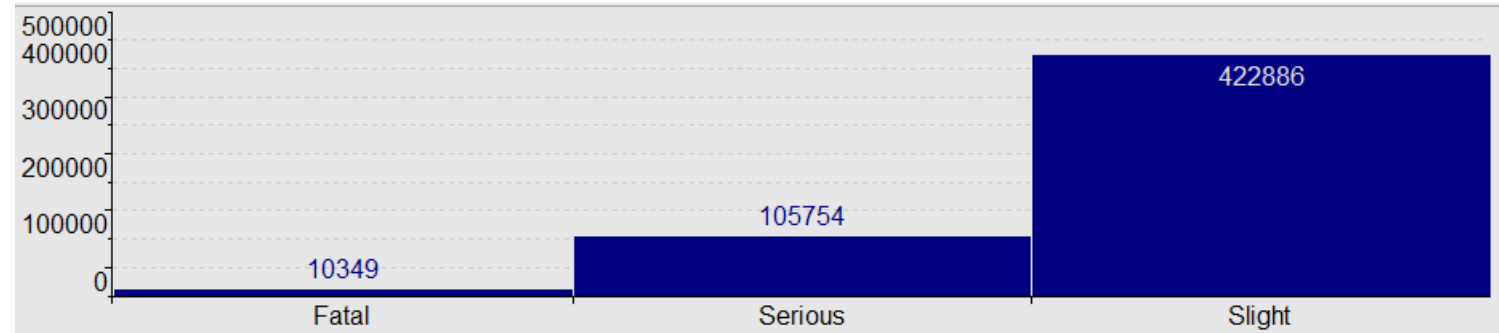


Weather_Conditions

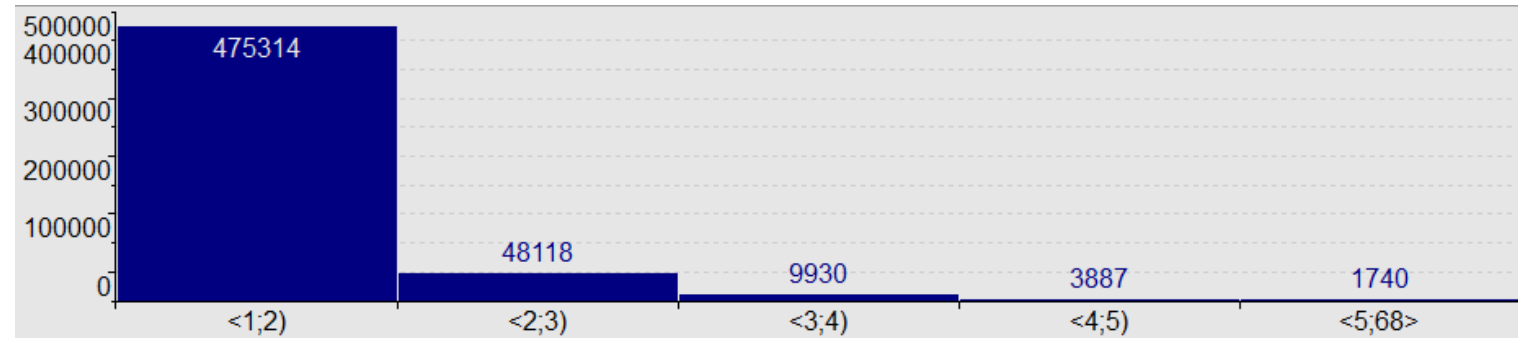


Data o nehodě

Accident_Severity



Number_of_Casualties
(nevyužito, inspirace
pro další analýzy)



Zadání procedury SD4ft-Miner (1)

DEPENDENT

Driver Con, 0 - 2
» Age_Groups_of_Driver (subset), 1 - 1 R, pos
» Driver_Home_Area_Type (subset), 1 - 1 B, pos
Circumstances Con, 0 - 4
» Junction_Detail (subset), 1 - 1 B, pos
» Light_Conditions (subset), 1 - 1 B, pos
» Road_Surface_Conditions (subset), 1 - 1 B, pos
» Weather_Conditions (subset), 1 - 1 B, pos

Total length: 0 - 6

QUANTIFIERS

Type	Rel.	Value	Units
a (BASE) FirstSet	>=	500.00	Abs
a (BASE) SecondSet	>=	500.00	Abs
PIM RatioVal	>	2.00	Abs

Generation information
Status: Solved, 28 run(s)
Mode: Standard

SUCCEDENT

Severity Con, 1 - 1
» Accident_Severity (subset), 1 - 1 B, pos

Total length: 1

CONDITION

Default Partial Cedent Con, 0 - 5

Total length: 0

Task parameters

Verification mode: The comparison of the first set versus the second set
Sets overlapping: Sets must differ in all rows (i.e. not overlapping sets)
Maximal number of hypotheses: 1000

(1) FIRST SET

Sex Con, 1 - 1
» Sex_of_Driver (subset), 1 B, pos

Total length: 1

(2) SECOND SET

Sex Con, 1 - 1
» Sex_of_Driver#2 (subset), 1 B, pos

Total length: 1

Automaticky generované dvojice podmnožin

Zadání procedury SD4ft-Miner (2)

QUANTIFIERS			
Type	Rel.	Value	Units
a (BASE) FirstSet	>=	500.00	Abs
a (BASE) SecondSet	>=	500.00	Abs
PIM RatioVal	>	2.00	Abs

Minimum pro a-frekvence pro první i druhé pravidlo

SD4ft Statistical quantifier settings

Interest measure type: **p-Implication**
Relation: Greater than
Threshold value: 2

Operation mode: Ratio of interest-measures

Test applied to the ratio of interest-measures each frequency table

$$\frac{\text{Konfidence Driver}(?) \wedge \text{Circumstancies}(?) \approx \text{Severity}(?) / \text{Sex}(?)}{\text{Konfidence Driver}(?) \wedge \text{Circumstancies}(?) \approx \text{Severity}(?) / \text{Sex}(?)} > 2$$

Výstup procedury SD4ft-Miner

Task run
Start: 12.10.2019 12:05:00 Total time: 0h 0m 10s
Number of verifications: 19116
Number of hypotheses: 20 Mode: Standard

Add

Každé z 20 vystupujících pravidel odpovídá situaci kdy

Je pro některé řidiče za nějakých okolností četnost poruch s nějakou závažností víc než dvakrát větší / menší pro muže než pro ženy.

V tomto případě je to vždy pro muže a týká se to závažnosti Fatal

Actual group of hypotheses: All hypotheses

Hypotheses in group: 20 Shown hypotheses: 20

Highlight

Nr. Id R-Conf 1:Conf 2:Conf Hypothesis

Nr.	Id	R-Conf	1:Conf	2:Conf	Hypothesis
1	19	2.300	0.026	0.011	Driver_Home_Area_Type(Urban area) & Junction_Detail(Not at junction or within 20 metres) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
2	20	2.239	0.022	0.010	Driver_Home_Area_Type(Urban area) & Road_Surface_Conditions(Dry) & Weather_Conditions(Fine no high winds) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
3	10	2.235	0.029	0.013	Junction_Detail(Not at junction or within 20 metres) & Weather_Conditions(Fine no high winds) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
4	18	2.235	0.022	0.010	Driver_Home_Area_Type(Urban area) & Road_Surface_Conditions(Dry) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
5	6	2.234	0.028	0.013	Junction_Detail(Not at junction or within 20 metres) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
6	11	2.223	0.020	0.009	Junction_Detail(Not at junction or within 20 metres) & Light_Conditions(Daylight) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
7	16	2.204	0.023	0.011	Driver_Home_Area_Type(Urban area) & Weather_Conditions(Fine no high winds) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
8	15	2.179	0.022	0.010	Driver_Home_Area_Type(Urban area) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
9	13	2.171	0.029	0.013	Junction_Detail(Not at junction or within 20 metres) & Road_Surface_Conditions(Dry) & Weather_Conditions(Fine no high winds) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
10	12	2.167	0.029	0.013	Junction_Detail(Not at junction or within 20 metres) & Road_Surface_Conditions(Dry) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
11	14	2.133	0.025	0.012	Road_Surface_Conditions(Wet or damp) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
12	2	2.128	0.025	0.012	Weather_Conditions(Fine no high winds) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
13	1	2.120	0.024	0.011	(empty) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
14	5	2.113	0.024	0.011	Road_Surface_Conditions(Dry) & Weather_Conditions(Fine no high winds) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
15	17	2.104	0.017	0.008	Driver_Home_Area_Type(Urban area) & Light_Conditions(Daylight) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
16	4	2.102	0.023	0.011	Road_Surface_Conditions(Dry) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
17	3	2.083	0.017	0.008	Light_Conditions(Daylight) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
18	9	2.052	0.018	0.009	Light_Conditions(Daylight) & Road_Surface_Conditions(Dry) & Weather_Conditions(Fine no high winds) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)
19	8	2.047	0.018	0.009	Light_Conditions(Daylight) & Road_Surface_Conditions(Dry) >>< Accident_Severity(Fatal) : Sex(Male) × Sex(Female)

Výstup procedury SD4ft-Miner - detail prvního SD4ft-pravidla

$\varphi \approx \psi$: Driver_Home_Area(Urban) \wedge Junction_Detail(Not_at_junction-or_within_20_metres) \approx Severity(Fatal)

Antecedent: Driver_Home_Area_Type(Urban area) & Junction_Detail(Not at junction or within 20 metres)

Succedent: Accident_Severity(Fatal)

First set: Sex(Male)

Second set: Sex(Female)

Condition: (empty)

TEXT | DATA | FIRST SET | SECOND SET | F+S SET | DIFF ABS | DIFF REL

$\varphi \approx \psi$ / Sex(Male)

$\varphi \approx \psi$ / Sex(Female)

	$\varphi \approx \psi$ / Sex(Male)		$\varphi \approx \psi$ / Sex(Female)	
	Succedent	\neg Succedent	Succedent	\neg Succedent
Antecedent	3296	121704	545	46988
\neg Antecedent	5318	234948	982	88755

$$\text{Konfidence } \varphi \approx \psi / \text{Sex(Male)} = \frac{3296}{3296+121704} = 0.0264$$

$$\text{Konfidence } \varphi \approx \psi / \text{Sex(Female)} = \frac{545}{545+46988} = 0.0115$$

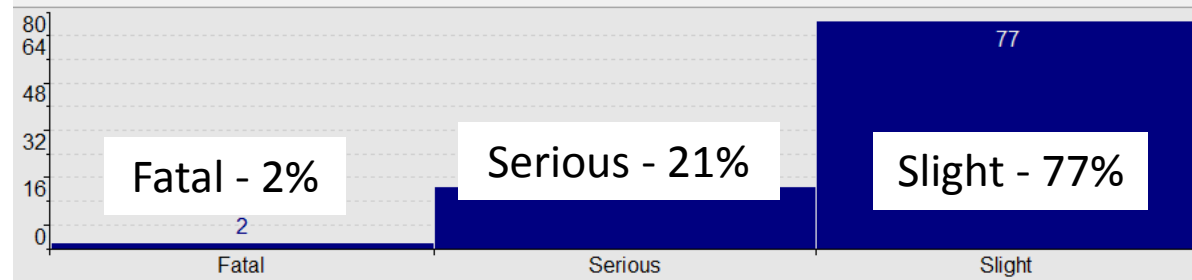
$$\frac{\text{Konfidence } \varphi \approx \psi / \text{Sex(Male)}}{\text{Konfidence } \varphi \approx \psi / \text{Sex(Female)}} = \frac{0.0264}{0.0115} = 2.3$$

Procedura SD4ft-Miner - využití poměru konfidencí

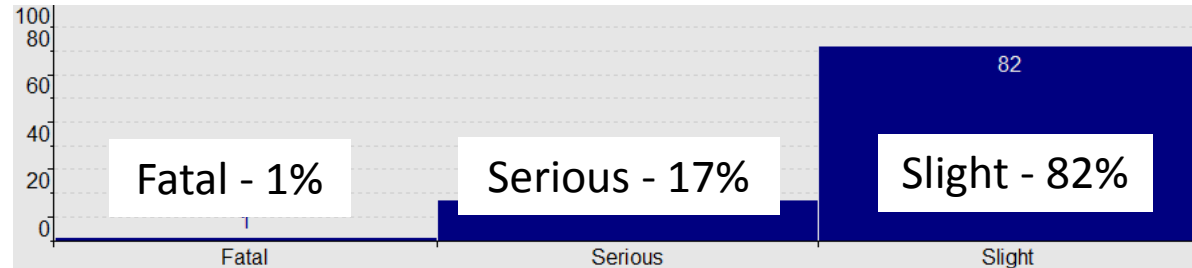
- Data Accidents
- Vysoký poměr konfidencí
- Poměr konfidencí přibližně jedna

Poměr konfidencí přibližně jedna - inspirace

Procentuální podíl hodnot atributu Accident_Severity pro řidiče - muže:



Procentuální podíl hodnot atributu Accident_Severity pro řidiče - ženy:



Otázka: Je pro některé řidiče za nějakých okolností relativní četnost poruch s nějakou závažností téměř stejná pro muže i pro ženy?

$$\frac{\text{Konfidence } \text{Driver}(?) \wedge \text{Circumstancies}(?) \approx \text{Severity}(?) / \text{Sex}(?)}{\text{Konfidence } \text{Driver}(?) \wedge \text{Circumstancies}(?) \approx \text{Severity}(?) / \text{Sex}(?)} \approx 1$$

Zadání procedury SD4ft-Miner (1)

ANTECEDENT

Driver Con, 0 - 2
» Age_Groups_of_Driver (subset), 1 - 1 R, pos
» Driver_Home_Area_Type (subset), 1 - 1 B, pos
Circumstances Con, 0 - 4
» Junction_Detail (subset), 1 - 1 B, pos
» Light_Conditions (subset), 1 - 1 B, pos
» Road_Surface_Conditions (subset), 1 - 1 B, pos
» Weather_Conditions (subset), 1 - 1 B, pos

Total length: 0 - 6

QUANTIFIERS

Type	Rel.	Value	Units
a (BASE) FirstSet	>=	500.00	Abs
a (BASE) SecondSet	>=	500.00	Abs
PIM RatioVal	>=	1.00	Abs
PIM RatioVal	<	1.00	Abs

Generation information
Status: Solved, 30 run(s)
Mode: Standard

SUCCEDENT

Severity Con, 1 - 1
» Accident_Severity (subset), 1 - 1 B, pos

Total length: 1

CONDITION

Default Partial Cedent Con, 0 - 5

Total length: 0

(1) FIRST SET

Sex Con, 1 - 1
» Sex_of_Driver (subset), 1 B, pos

Total length: 1

(2) SECOND SET

Sex Con, 1 - 1
» Sex_of_Driver#2 (subset), 1 B, pos

Total length: 1

Task parameters

Verification mode: The comparison of the first set versus the second set
Sets overlapping: Sets must differ in all rows (i.e. not overlapping sets)
Maximal number of hypotheses: 1000

Driver(?)

Circumstances(?)

Severity(?)

Sex(?)

Sex(?)

Automaticky generované dvojice podmnožin

Zadání procedury SD4ft-Miner (2)

QUANTIFIERS			
Type	Rel.	Value	Units
a (BASE) FirstSet	>=	500.00	Abs
a (BASE) SecondSet	>=	500.00	Abs
PIM RatioVal	>=	1.00	Abs
PIM RatioVal	<	1.00	Abs

Minimum pro a-frekvence pro první i druhé pravidlo

SD4ft Statistical quantifier settings

Interest measure type: p-Implication

Relation: Greater than or equal

a/(a+b) >= p ... at least 100*p [%] of objects satisfying A satisfy also S

Threshold value: 1

Operation mode: Ratio of interest-measures

SD4ft Statistical quantifier settings

Interest measure type: p-Implication

Relation: Less than

a/(a+b) >= p ... at least 100*p [%] of objects satisfying A satisfy also S

Threshold value: 1.001

Operation mode: Ratio of interest-measures

$$1 \leq \frac{\text{Konfidence Driver}(?) \wedge \text{Circumstancies}(?) \approx \text{Severity}(?) / \text{Sex}(?)}{\text{Konfidence Driver}(?) \wedge \text{Circumstancies}(?) \approx \text{Severity}(?) / \text{Sex}(?)} < 1.001$$

Výstup procedury SD4ft-Miner

Task run

Start: 12.10.2019 17:55:12

Total time: 0h 0m 10s

Number of verifications: 19116

Number of hypotheses: 4

Mode: Standard

Add group

Del group

Edit group

Actual group of hypotheses: All hypotheses

Hypotheses in group: 4

Shown hypotheses: 4

Highlighted: 0

Delete hypotheses

Nr. Id R-Conf 1:Conf 2:Conf Hypothesis

1	4	1.000	0.803	0.803	Age_Driver(56 to 65) & Driver_Home_Area_Type(Urban area) & Junction_Detail(T or staggered junction) & Light_Conditions(Daylight) >=< Accident_Se
2	2	1.001	0.756	0.756	Age_Driver(76 to 85) & Driver_Home_Area_Type(Urban area) & Light_Conditions(Daylight) >=< Accident_Severity(Slight) : Sex(Male) × Sex(Female)
3	3	1.001	0.191	0.191	Junction_Detail(Crossroads) & Light_Conditions(Daylight) & Road_Surface_Conditions(Dry) >=< Accident_Severity(Serious) : Sex(Female) × Sex(Male)
4	1	1.001	0.792	0.791	Age_Driver(26 to 35) & Driver_Home_Area_Type(Urban area) & Junction_Detail(Crossroads) & Light_Conditions(Daylight) & Road_Surface_Condition

Každé ze čtyř vystupujících pravidel odpovídá situaci kdy

Pro některé řidiče je za nějakých okolností četnost poruch s nějakou závažností téměř stejná pro muže i pro ženy.

Výstup procedury SD4ft-Miner - detail prvního SD4ft-pravidla

ϕ : Age_Driver(56 to 65) \wedge Driver_Home_Area(Urban) \wedge Junction_Detail(T or staggered junction) \wedge Light_Conditions(Daylight)
 ψ : Accident_Severity(Slight)

First set: Sex(Female)

Second set: Sex(Male)

Condition: (empty)

TEXT | DATA | FIRST SET | SECOND SET | F+S SET | DIFF ABS | DIFF REL

$\phi \approx \psi$ / Sex(Female)

$\phi \approx \psi$ / Sex(Male)

	$\phi \approx \psi$ / Sex(Female)		$\phi \approx \psi$ / Sex(Male)	
	Succedent	\neg Succedent	Succedent	\neg Succedent
Antecedent	1275	313	4111	1010
\neg Antecedent	111492	24190	275529	84616

$$\text{Konfidence } \phi \approx \psi / \text{Sex(Female)} = \frac{1275}{1275+313} = 0.8029$$

$$\text{Konfidence } \phi \approx \psi / \text{Sex(Male)} = \frac{4111}{4111+1010} = 0.8028$$

$$\frac{\text{Konfidence } \phi \approx \psi / \text{Sex(Male)}}{\text{Konfidence } \phi \approx \psi / \text{Sex(Male)}} = \frac{0.8029}{0.8028} = 1.0002$$