

1. БАЕСОВИ МРЕЖИ

Платежни измами

Дефиниција на проблемот:

Секоја година фирмите кои издаваат платежни картички трпат милионски загуби поради измами поврзани со изгубени или украдени картички. Задачата е да се моделира систем со ВИ кој ќе детектира измами при плаќањата. Интуитивно, сопствениците на платежни картички купуваат според одредена повторувачка шема, па постои поголема веројатност дека станува збор за измама кога оваа шема е нарушена. Според претходното искуство, добиени се следниве статистички показатели:

- 1,2% од трансакциите се измамнички кога корисниците патуваат во друго место, за разлика од само 0,4% кога не патуваат.
- Во просек, 5% од сите трансакции се случуваат кога корисниците патуваат.
- 10% од измамничките трансакции се случуваат во странство додека патуваат.
- Кога корисниците не патуваат, 10% од измамничките трансакции се направени од странство.
- Од сите трансакции, само 1% од правилните трансакции се направени во странство.
- Кога патуваат, 35% од сите трансакции се од странство без разлика на законитоста на трансакциите.

Многу поголема веројатност за измамнички трансакции е при плаќањата преку Интернет, посебно за оние корисници кои немаат компјутер.

- 90% од населението има некаков вид компјутер и таквите корисници 4% од нивните правилни трансакции ги прават преку Интернет, додека процентот на измамнички трансакции кај таквите корисници преку Интернет е 2%.
- Кај тие што немаат компјутер, само 0,01% од нивните правилни трансакции се направени преку Интернет, но тој број расте на 1,5% кај нив за измамнички трансакции преку Интернет.
- Дали некој има или нема компјутер се утврдува според купувањето на компјутерска опрема, па така 7% од тие што имаат компјутер купиле (со нивната картичка) некаква компјутерска опрема, за разлика од само 0,02% од оние кои немаат никаков компјутер.

За дадениот проблем кој треба да се моделира како Баесова мрежа потребно е да одговорите на следните прашања/задачи поделени по делови:

(а) Да се претстават случајните променливи со кои го моделирате проблемот користејќи Баесова мрежа.

(ПОМОШ: Обрнете внимание на зборовите со задебелени букви во описот погоре:

Трансакциите можат да бидат: измамнички или правилни.

Трансакциите можат да бидат направени преку Интернет или не.

Корисниците можат да имаат или да немаат компјутер.

Корисниците можат да **патуваат** или да не патуваат.

Кога патуваат, може да патуваат во **странство** или во земјава.

И може да купиле **компјутерска опрема** или да не купиле.)

(б) Да се даде графички приказ на Баесовата мрежа!

(в) Колку вкупно параметри (поединечни вредности на веројатностите) се потребни за да се дефинираат условните веројатносни распределби (во табелите покрај секој јазол) за оваа Баесова мрежа?

(г) Да се прикажат маргиналните и условни веројатносни табели користејќи ги податоците од описот на проблемот. За вредностите за кои немате податоци изберете вредности според ваш избор.

(д) Напишете го изразот за тотална здружена веројатност за случајните променливи кои се дефинирани во Баесовата мрежа. Користете го верижното правило според подредување.

(ѓ) Наведете ги сите независности (апсолутни и условни) помеѓу променливите кои произлегуваат од структурата на Баесовата мрежа. Образложете!

(е) Напишете го изразот за веројатност дека трансакцијата е **измамничка**, ако се знае дека корисникот во тој момент **не** бил во **странство**. Да се пресмета нумеричката вредност на оваа веројатност, користејќи ги вредностите од веројатносните табели.

(ж) Напишете го изразот за веројатност дека трансакцијата е **измамничка**, ако се знае дека корисникот купил неодамна **компјутерска опрема**. Да се пресмета нумеричката вредност на оваа веројатност, користејќи ги вредностите од веројатносните табели.

2. МАШИНСКО УЧЕЊЕ

Полицијата во некој град сака да автоматизира дел од своето работење преку воведување на машинско учење. Полицијата сака да може автоматски да одлучува дали постои опасност за животот на нејзините вработени врз основа на карактеристиките на случаите на терен на кои се упатени и тоа: дали теренскиот случај е во град или не, дали теренскиот случај се одвива преку ден или не и дали теренскиот случај се одвива поради повик на телефонот за итни случаи. За таа цел ги имате следните податоци:

#случај	Град	Ден/Ноќ	Итен повик	Опасност?
1	не	ноќ	не	Висока
2	не	ноќ	да	Висока
3	не	ноќ	да	Ниска
4	не	ден	не	Ниска
5	не	ден	не	Ниска
6	да	ноќ	не	Висока
7	да	ноќ	не	Висока
8	да	ден	да	Ниска
9	да	ден	да	Ниска
10	да	ден	да	Висока

За секој од алгоритмите наведени во продолжение дајте одговор на соодветните барања користејќи ги податоците дадени во табелата како множество на обука.

А. Наивен Баесов класификатор

Треба да истренирате наивен Баесов класификатор за предвидување на веројатноста за нивото на опасност по животот на полицајците со користење на Лапласово порамнување со $k=2$ виртуелни примероци.

- (а) Нацртајте го моделот на наивниот Баесов класификатор и прикажете ги сите негови параметри.
- (б) Како ќе биде класифициран случај кој не е во град, се одвива дење и имало итен повик за истиот?
- (в) Колкава е веројатноста за случај кој е во град во текот на ноќта да има високо ниво на опасност по животот на полицајците?

Б. Перцептрон

Сакате да употребите перцептрон за предвидување на нивото на опасност по животот на полицајците.

- (а) Предложете соодветен модел на перцептрон. Поставете иницијални вредности за параметрите на перцептронот различни од 0 (нула). Каков е ефектот на параметри еднакви на 0 (нула) во однос на тоа како перцептронот предвидува соодветна класа? Каков ќе биде ефектот на вредностите кои вие ги поставивте?
- (б) Покажете една епоха од учењето на перцептронот.

(в) Дали постои гаранција дека алгоритмот за учење на перцептронот ќе конвергира?

В. Дрво на одлучување

На крај сакате да изградите дрво на одлучување, при што како критериум за избор на најдобар атрибут за поделба користите информациска добивка.

(а) Покажете ја целосната постапка (вклучително и сите пресметки) за избор на најдобар атрибут за поделба во коренот на дрвото на одлучување.

(б) Покажете го целосното дрво на одлучување без да ги покажувате пресметките. За секој лист во дрвото на одлучување дајте образложение зошто ја предвидува класата која сте ја доделиле.