

Zajęcie 7. Podejmowanie decyzji w warunkach ryzyko. Użycie drzewa decyzji

Abstract

Celem jest podejmowanie decyzji za pomocą indukcji drzewa decyzji na podstawie algorytmu C5.0

1. Wprowadzenie

Drzewa decyzyjne stanowią model decyzyjny, w którym w uporządkowany sposób przedstawia się hierarchiczne ciągi działań (w pełni zależnych od decydenta) i zdarzeń (niezależnych od decydenta, czasami o charakterze losowym). Graficzne przedstawienie w postaci drzewa decyzyjnego ułatwia analizę wszystkich elementów sytuacji istotnych przy podejmowaniu decyzji. W efekcie możliwe staje się określenie wariantów decyzyjnych i ich konsekwencji. W modelu tym nie występują tu w jawnej postaci warunki sztywne i elastyczne, są one uwzględniane w trakcie budowy drzewa. Dodatkowe podanie prawdopodobieństw i kosztów poszczególnych wariantów decyzyjnych prowadzi do zwiększenia racjonalności optymalizacyjnej poprzez maksymalizację funkcji użyteczności. Celem stosowania modelu w postaci drzewa decyzyjnego jest uproszczenie oceny sytuacji decyzyjnej, model ten pozwala na jednoczesną analizę wielu wariantów decyzyjnych i kryteriów ich oceny. Model taki jest użyteczny, o ile drzewo nie staje się zbyt obszerne (nie mieści się na kartce lub ekranie). Z wykorzystaniem drzew decyzyjnych może być prowadzona analiza wielowariantowa (what-if analysis), a poprzez implementację programową możliwe jest zastosowanie tego modelu w komputerowych systemach wspomagania decyzji.

Przykład. Przykład typowego drzewa decyzyjnego (2-poziomowego) przedstawiono na Rys. 1.

Rozważana jest tutaj sytuacja związana z ubezpieczeniem mieszkania, przy założeniu kosztów polisy w wysokości 3% oraz wkładu własnego w wysokości 2% wartości mieszkania. Możliwym zdarzeniom (brak kradzieży,

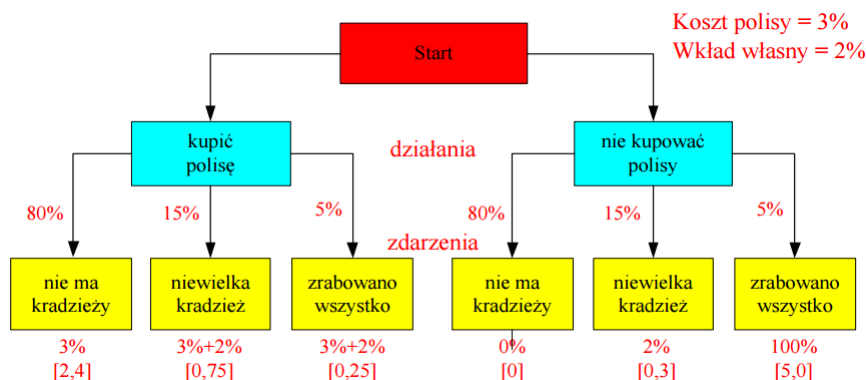


Figure 1: Drzewo decyzyjne dot. decyzji o zakupie polisy na ubezpieczenie mieszkania

kradzież niewielka – nie przekraczająca wkładu własnego oraz kradzież pełna) przypisano prawdopodobieństwa ich wystąpienia (odpowiednio 80, 15 i 5%). Zarówno wydatki związane z zakupem polisy, z ponoszeniem wkładu własnego, jak i rekompensatą za skradzione wyposażenie mieszkania (w przypadku rezygnacji z zakupu polisy) traktowane są jako strata, którą należy zminimalizować.

Wartość oczekiwaną straty związanej z daną decyzją można obliczyć z zależności (1.2) wprowadzonej przy okazji omawiania strategii skalania prawdopodobieństw i użyteczności, przy czym tutaj użytecznością (negatywną – strata) będzie koszt poniesiony przy danym wariancie decyzyjnym dla poszczególnych zdarzeń. Wartości iloczynów $\pi_i(\text{sk})$ $u_i(\text{sk})$ umieszczono w nawiasach kwadratowych pod zdarzeniami na Rys. 1. Oczekiwana strata dla poszczególnych decyzji wynosi

$$SPU(d_i) = \begin{cases} 3.4, & \text{dla decyzji kupować polisę} \\ 5.3, & \text{dla decyzji nie kupować polisy} \end{cases}$$

a zatem właściwą decyzją będzie zakup polisy ubezpieczeniowej.

2. Warianty Zadania

Zadanie dotyczy prognozowania oceny klientów (w skali 5-punktowej, Error < 5%) urządzeń RTV AGD. Używając metody indukcji drzewa decyzyjnego C5.0 opracować plik w języku R z wykorzystaniem paczki C50. Sprawozdania w postaci pliku R, obrazu drzewa decyzyjnego oraz wyników z konsoli

(dowolny plik tekstowy) zachować w zdalnym repozytorium (np Github) link na który wysłać w mailu z tematem **SMPD_7_Gr_numer_grupy** na adres mailowy **vmartsenyuk@ath.bielsko.pl**

1. Smartfon Samsung z systemem Android 6, ośmiordzieniowym procesorem. Uwzględniamy następujące dane: wyświetlacz, pamięć RAM, pamięć wbudowana, aparat foto. Komunikacja (wifi itp) nie ma znaczenia. Dane (>10 smartfonów) pobrać ze strony <http://www.euro.com.pl>
2. Macbook Apple. Uwzględniamy następujące dane: ekran, pamięć RAM, dysk i ewentualnie jakieś inne parametry. Dane (>10 Macbooków) pobrać ze strony <http://www.euro.com.pl>
3. Aparat z wymienną optyką. Uwzględniamy następujące dane: rozdzielczość, zakres czułości, obiektyw w zestawie itp. Dane (12 aparatów) pobrać ze strony <http://www.euro.com.pl>
4. Lodówka. Uwzględniamy następujące dane: pojemność użytkową chłodziarki, pojemność użytkową zamrażarki oraz jakieś inne parametry. Dane (15 lodówek) pobrać ze strony <http://www.euro.com.pl>
5. Lodówka turystyczna typu elektrycznego. Uwzględniamy następujące dane: pojemność, wagę itp. Dane (20 lodówek) pobrać ze strony <http://www.euro.com.pl>
6. Pralka ze sposobem załadunku od przodu. Uwzględniamy następujące dane: pojemność, maksymalna prędkość wirowania itp. Dane (20 pralek) pobrać ze strony <http://www.euro.com.pl>
7. Tablet iPad z systemem iOS 9, 2-rdzeniowym procesorem, modemem LTE. Uwzględniamy następujące dane: wyświetlacz, pamięć RAM i pamięć wbudowaną itp. Dane (10 iPadów) pobrać ze strony

<http://www.euro.com.pl>

8. Monitor LED. Uwzględniamy następujące dane: matrycę (pikseli), jasność ekranu, czas reakcji matrycy itp. Dane (>10 monitorów) pobrać ze strony
<http://www.euro.com.pl>
9. Smartfon Microsoft. Uwzględniamy następujące dane: wyświetlacz, pamięć RAM i pamięć wbudowaną, aparat foto itp. Dane (>10 smartfonów) pobrać ze strony
<http://www.euro.com.pl>
10. Smartfon Lenovo z systemem Android 6, ośmiokdzeniowym procesorem. Uwzględniamy następujące dane: wyświetlacz, pamięć RAM i pamięć wbudowaną, aparat foto, komunikacja (wifi itp). Dane (>10 smartfonów) pobrać ze strony
<http://www.euro.com.pl>
11. Smartfon z dual sim Huawei z systemem Android 6, ośmiokdzeniowym procesorem. Uwzględniamy następujące dane: wyświetlacz, pamięć RAM i pamięć wbudowaną, aparat foto, komunikacja (wifi itp). Dane (>10 smartfonów) pobrać ze strony
<http://www.euro.com.pl>
12. Smartfon iPhone z czterodzeniowym procesorem architektury 64-bit. Uwzględniamy następujące dane: wyświetlacz, pamięć RAM i pamięć wbudowaną, aparat foto, cena, komunikacja (wifi itp). Dane (>10 smartfonów) pobrać ze strony
<http://www.euro.com.pl>

References

- https://pl.wikipedia.org/wiki/Drzewo_decyzyjne
<https://www.dezyre.com/data-science-in-r-programming-tutorial/decision-tree-tutorial>
<https://static1.squarespace.com/static/51156277e4b0b8b2ffe11c00/t/51e67b45e4b0e6c130fb4d5>
<http://en.proft.me/2016/11/9/classification-using-decision-trees-r/>

Youtube: <https://www.youtube.com/watch?v=5NquIfQxpxk>