

Prof. dr hab. Andrzej Sokołowski
Uniwersytet Ekonomiczny w Krakowie

DZIAŁ
Kolegium Analiz Ekonomicznych

dr hab. Joanna Piębańiak, prof. SGH

9.11.17

Recenzja pracy doktorskiej mgra Damiana Przekopa pt. „Algorytmiczna budowa predyktorów w modelowaniu zdarzeń rzadkich na przykładzie wykrywania oszust kredytowych”

Tytuł dysertacji dobrze oddaje zamiar Autora i zawartość pracy. Najciekawszym elementem są dwa algorytmy zaproponowane do budowy predyktorów, które następnie wykorzystywane są do identyfikacji zdarzeń rzadkich. Te propozycje mogą być wykorzystywane w analizie różnych problemów, ale w pracy skoncentrowano się słusznie na wykrywaniu oszust kredytowych. Dobrym zabiegiem było podzielenie pracy na dwie części. Pierwsza ma charakter teoretyczny, a w drugiej Autor przedstawia własne propozycje oraz wyniki własnych badań. Podział pierwszej części na trzy rozdziały poświęcone kolejno: oszustwom kredytowym w sektorze bankowym, analitycznym metodom wykrywania nadużyć oraz obszarom wykrywania nadużyć oceniam jako trafny. Autorowi udało się dobrać właściwy poziom szczegółowości omawianych zagadnień – bez popadania z jednej strony w ogólniki, a z drugiej w podręcznikową drobiazgowość. Uważam za celowe przedstawienie nawet takich metod, które w drugiej części nie były potem stosowane (np. samoorganizujące się mapy Kohonena), bo pozwala to czytelnikowi zobaczyć szerszy obraz metod stosowanych w wykrywaniu oszust kredytowych.

Podział pracy na dwie części skutkował tym, że hipotezy badawcze sformułowano dopiero na 68 stronie rozprawy (nie licząc wstępu). Gorzej, że sformułowano one są dość pokrętnie. W hipotezie pierwszej jest napisane, że nowy algorytm służy do konstrukcji zmiennych. Zatem nie może on służyć do prognozowania. Do prognozowania można wykorzystać te zmienne. W hipotezie drugiej chodzi w zasadzie o to samo (czyli lepsze prognozowanie) przy pomocy algorytmu konstrukcji zmiennych (??) i dalej „opartego na koncepcji wykorzystania informacji mającej swe źródło w logice reguł decyzyjnych”. Po pierwsze użyto złego przypadku. Po drugie - wydaje mi się, że reguły decyzyjne są jakie są – jednakowe na wszystkich obiektów (!) i nie dostarczają żadnej informacji. Cel badania (s.62) też jest sformułowany niezręcznie. Celem badania nie jest poprawa jakości predykcji tylko zaproponowanie i przetestowanie takich metod (algorytmów), które pozwalają lepiej identyfikować kredyty zagrożone.

Najślabszą stroną rozprawy jest jej język – momentami straszny. Wszystkie te uchybienia wymieniam w uwagach szczegółowych. Autor wielokrotnie używa niewłaściwie takich wyrazów jak: dedykowany, sekcja, klasteryzacja, adresować, inicjalny. Niepotrzebnie mamy detekcję zamiast wykrywanie. Ta nieporadność językowa wprawia czytelnika w niepotrzebną irytację i utrudnia lekturę.

Z zagadnień merytorycznych mam następujące uwagi:

- Zbytнім skrótem myślowym jest uznawanie przez Autora percentyli za przedziały liczbowe, podczas gdy są to liczby.
- Wybór 5, 10, 15, 20 głównych składowych to decyzja trochę „z sufitu”. Należało przyjąć kryterium wartości własnej większej od 1 lub zadanego procentu wyjaśniania wariancji.
- Nieprecyzyjny jest opis zastosowania metody k-średnich (s.83). Co grupowano – zmienne czy obiekty? Prawdopodobnie zmienne, ale dlaczego „poddane one zostały działaniu algorytmu grupowania obserwacji (a więc może jednak grupowano obserwacje, a nie zmienne- AS) opartego na koncepcji k-średnich” Czy więc zastosowano metodę k-średnich, czy jakąś inną „opartą na jej koncepcji” (np. k-medoidów)?
- W jaki sposób przeprowadzono agregację rozkładów dwóch zmiennych ciągłych przy tworzeniu zbioru D (s.94)
- Wypadało podać listę zmiennych wyjściowych wykorzystywanych w badaniu (191 zmiennych ciągłych i 22 dyskretnych – choć tu mam podejrzenie, że nie chodzi o zmienne dyskretne, tylko o jakościowe)
- Szanuję wybór miary Lift do oceny jakości wykrywania oszust, choć chciałbym też zobaczyć jak kształtuje się *dodatnia i ujemna zdolność predykcyjna, oraz czułość i swoistość*.

Wysoko cenię pomysłowość Autora, która stanowi główną zaletę ocenianej rozprawy. Proponowane algorytmy zostały opisane mało precyzyjnie, ale mam wrażenie, że udało mi się odkryć ich ideę. Pierwszy algorytm trochę przypomina mi koncepcję lasów losowych, ale trzeba go uznać za oryginalny w kontekście całości problemu, czyli wykorzystania zmiennych oryginalnych i tych tworzonych nowym algorytmem. Bardzo ciekawy jest drugi zaproponowany algorytm, łączący wykorzystanie informacji jakościowej (identyfikacja grup odniesienia) z rozkładami zmiennych losowych ciągłych. Słusznie ten algorytm jest szeroko obecny w modelach na s.91-94 (w tym bardzo dobry Rysunek 3). Szanuję wybór LASSO jako metody doboru zmiennych, choć mam do tej metody pewne zastrzeżenia.

Mocną stroną pracy są jej dwa ostatnie punkty 5.3 oraz 5.4. Bardzo przekonująco wygląda siedem schematów eksperymentu. Badania przeprowadzone w ich ramach pozwoliły dobrze ocenić stabilność ze względu na dobór próby, reedukację liczby zmiennych oraz definicję funkcji celu. Interpretacja wyników jest poprawna. Ciekawe są pewne proste próby wyjaśnienia rezultatów. Dobrze, że obok krzywych Lift – często słabo rozróżnialnych – pojawiły się tabele z wartościami tej miary dla przedziału liczbowego od $Q_{0,95}$ do maksimum.

Kładąc na szalę słaby poziom językowy pracy i jej walory merytoryczne stwierdzam, że te ostatnie zdecydowanie przeważają. Literacki poziom wypowiedzi można poprawić (wyzbywając się choćby nadmiernej fascynacji pseudonaukowymi naleciałościami z angielszczyzny), natomiast jak ktoś nie potrafi proponować nowych rozwiązań z wykorzystaniem dotychczasowego dorobku metodologicznego – to jest stracony dla nauki. W tej rozprawie Autor pokazał zdolność proponowania własnych oryginalnych rozwiązań, wbudowywania ich w szerszy kontekst wykrywania zagrożeń oszustwami, a wreszcie dobrze przeprowadził proces weryfikacji celu pracy, czyli zaproponowania takich rozwiązań, które pozwalają na lepsze prognozowanie oszustw kredytowych.

W związku z powyższym stwierdzam, że moim zdaniem, praca spełnia wymogi stawiane rozprawom doktorskim i wnoszę o dopuszczenie Autora do publicznej obrony.

Uwagi szczegółowe

Strona (s) 6, wiersz (w) 4, od góry (g) – Dziwnie brzmi stwierdzenie, że „*Badanie ... zostało oparte na zjawisku nadużyć ...*”

s.8, w.2, od dołu (d) – Dziwnie brzmią „*oszustwa popełniane w kanale*”, nawet jeśli jest to kanał (??) bankowości elektronicznej

s.8, w.2, d – Zdanie „*Nietypowość oraz rzadkość wymienionych zjawisk (chodzi o te oszustwa w kanale – AS) czyni je zagadnieniami o zbliżonej problematyce do poruszanego*”. Co lub kto jest tym „poruszonym”? (rodzaju męskiego, a zagadnienia i zjawiska są rodzaju nijakiego). Lepszy byłby szyk „o problematyce zbliżonej”

s.9, w.7, g – Lepiej „*przeciwwyłudzeniowe*” niż „*antyfraudowe*”. W języku polskim nie ma wyrazu „*fraud*”. To samo: s.12, w.7, g

s.10, w.17, d – Dyskusyjne jest twierdzenie o „*niemożności budowy wiarygodnych próbek sumulacyjnych ...*”. Być może Autorowi chodziło o trudności w budowie wiarygodnych scenariuszy alternatywnych. Jeżeli jest scenariusz, to pozyskanie próbek symulacyjnych jest banalne.

s.11, w.13, g – W Słowniku Języka Polskiego wyraz „*detekcja*” tłumaczony jest jako „*wykrywanie promieniowania jonizującego, cząstek elementarnych lub substancji chemicznych*”. Dlaczego Autor nie używa po prostu wyrazu „*wykrywanie*”. To samo jest na stronach: s.11, w.4, d; s.40, w.1, g; s.54, w.1,d; s.58, w.6, g; s.66, w.9, d; s.69, w.3, g; s.70, w.11, g; s.72, w.5, g; s.75, w.11, g; s.77, w.5, g

s.11, w.16, d – „*Outlier*” to nie jest obserwacja „*nietykowa*”, tylko „*odstająca*”. To samo: s.31, w.8, g

s.11, w.11, d – Zdecydowanie nie zgadzam się z twierdzeniem, że „*oszustwa kredytowe należy postrzegać jako obserwacje nietykowe*”. Zachowanie kredytobiorcy to cecha jakościowa, która ma dwa warianty: „*nie oszukuje*”/„*oszukuje*”

s.12, w.4, g – Co to „*wymiar dysertacji*” i dlaczego sam Autor twierdzi, że jest on „*unikalny*”

s.12, w.7, g – Autor zdecydował, że w pracy nie będzie badań symulacyjnych. Ja ten wybór szanuję, ale to recenzenci mają ocenić czy jest on „*uzasadniony*”, a nie Autor.

s.13, w.9, d – W zdaniu zaczynającym się od „*Wymaga ..*” dowiadujemy się, że z ryzykiem można się obchodzić, oraz można do niego podchodzić. Jestem ciekaw jak?

s.21, w.13, g – Językoznawcy wielokrotnie zwracali uwagę na to, że „*dedykować*” można jedynie utwór artystyczny. Informatycy są na to głusi i stąd mamy na przykład „*oprogramowanie dedykowane sferze finansów*”, a to jest po prostu „*oprogramowanie dla*

finansistów” i nikomu ono nie jest dedykowane. To samo: s.62, w.3, d; s.70, w.4, g; s.70, w.14, g; s.80, w.2, d; s.81, w.5, d; s.87, w.6, d; s.89, w.9, d

s.35, w.14, g – W opisie metody AID pomyłono dwa testy. Statystyka t występuje w teście Studenta (porównującym dwie grupy), a w analizie wariancji (porównującej wiele grup, choć można i dwie) występuje statystyka F.

s.36, w.3, g – Słownik języka polskiego tak wyjaśnia znaczenie słowa „adresować”: 1. «umieszczać adres na przesyłce», 2. «kierować, stosować lub odnosić coś do kogoś». Nie rozumiem więc jak można „adresować ten problem poprzez generowanie zestawu modeli”. Cambridge Dictionary podaje jedno z wyjaśnień czasownika „to address”: „to deal with a problem” i o to pewnie Autorowi chodziło. Zwracam uwagę, że poprawne tłumaczenie polega na znajdowaniu właściwego słowa w języku polskim, a nie czytanie angielskiego słowa po polsku (bo to jest wymyślony przez mnie język EngPol). To samo w następujących miejscach: s.36, w.2, d; s.64, w.3, g; s.67, w.7, g; s.67, w.4, d; s.70, w.11, g; s.81, w.1, d; s.87, w.12 (dwukrotnie) i 14, g; s.90, w.3, g

s.39, w.5, g – Co to są „silnie niezbilansowane problemy”. Słownik Języka Polskiego PWN podaje trzy znaczenia słowa „zbilansować”: 1. «sporządzić bilans», 2. «zrównoważyć pozycje w bilansie», 3. «zestawić i porównać coś z czymś». Które z nich ma tu zastosowanie?

s.42, w.4 oraz 16, g – Znowu język EngPol polegający na czytaniu po polsku wyrazów angielskich zamiast ich tłumaczenia. W analizie wielowymiarowej nie ma klastrów i klasteryzacji, natomiast są skupienia, skupiska, grupy, analiza skupień. To samo: s.46, w.6, g; s.47, w.5, g; s.49, w.5, 8, 9, d; s.55, w.5, g; s.56, w.3, d; s.69, w.7, g; Tytuł punktu 5.2.2; s.83, w.6 i 8, d; s.84, w.4, g; s.84-85, Tabele 4-10; s.85, w.6, d; s.86, w.8 i 10, g

s.43, w.12, g – Obserwacja odstająca nie charakteryzuje się rzadko występującą kombinacją zmiennych (bo wszystkie obserwacje opisane są tymi samymi zmiennymi), tylko wartości zmiennych.

s.63, w.1, g – Trudno zgodzić się z twierdzeniem, że literatura w zakresie modeli przeciwwyłudzeniowych jest uboga. Po wpisaniu w Google Scholar frazy „antifraud models” pokazuje się 9360 pozycji literatury.

s.64, w.7, g – Co to jest „próba przelosowana”? Czy jest to przeciwieństwo próby „niedolosowanej”?

s.64, w.15, g – Znowu Autor myli cechy z ich wartościami (lub wariantami). To samo: s.64, w.1, d

s.64, w.12, d – Na czym polega „budowanie należytej intuicji”?

s.64, w.2, d – Dlaczego przetłumaczono tylko połowę angielskiego wyrażenia tworząc potworek „socjoekonomiczny” zamiast „społeczno-ekonomiczny”?

s.65, w.4, d – Lepszy byłby szyk” ... przez algorytmy wprowadzone w niniejszej pracy”

s.66, w.14, g – Tragiczne zdanie zaczynające się od „W przypadku ...”. Słownik Języka Polskiego podaje jako pierwsze takie określenie wyrazu „selekcja”: «dobór osób lub rzeczy

przez eliminację tych, które nie spełniają określonych warunków», czyli eliminuje się „złe”, a „pozostawia dobre”. Oczywiście nie wybiera się „liczby wniosków” tylko „wnioski” i to nie ta liczba jest „obciążona najwyższym ryzykiem nadużycia”, tylko wnioski lub wnioskodawcy.

s.67, w.1 g, Słownik Języka Polskiego PWN podaje następujące znaczenia słowa „deficyt”:
1. «brak dostatecznej ilości pieniędzy w budżecie państwa, przedsiębiorstwa itp. na pokrycie zaplanowanych wydatków», 2. «niedostateczna ilość czegoś, niepokrywająca zapotrzebowania». A algorytm może mieć wady lub niedoskonałości, ale nie może mieć deficytu. To samo: s.70, w.10, g

s.69, w.6, g – Dlaczego Autor używa okropnej „centroidy klastra” (dlaczego centroida jest rodzaju żeńskiego?) na coś tak prostego jak środek grupy (skupienia)

s.69, w.14, g – Zły szyk zdania Ma by być „ ... lekarze generujący dla ubezpieczyciela najwyższe koszty ...”

s.69, w.4, d – Niepotrzebnie powtórzono definicję Hawkinsa

s.70, w.5, d – Zdanie „W ramach ...” pełne jest nieścisłości. W każdym podręczniku statystyki opisowe można przeczytać, że cechy to własności jednostek statystycznych (obiektów); cechy ilościowe są nazywane zmiennymi, a charakterystyki to miary położenia, zmienności, asymetrii itp. I to nie reguły są generowane przez drzewa decyzyjne, tylko akurat na odwrót – drzewa decyzyjne to nieliniowa sekwencja reguł. W związku z powyższym powoływane zdanie jest niezrozumiałe.

s.71, w.2, g – Niezgrabne sformułowanie „ ... oszczędność wykorzystywanych w modelowaniu stopni swobody.” Po pierwsze szyk zdania jest niewłaściwy, bo przecież nie modelujemy stopni swobody. One też nie są „wykorzystywane” w modelowaniu.

s.71, w.3, g – Słownik Języka Polskiego PWN podaje dwa znaczenia wyrazu „ekstrakcja”:
1. «usuwanie np. zęba lub soczewki oka», 2. «wyodrębnianie określonych składników mieszanin za pomocą rozpuszczalników». Co to jest „ekstrakcja reguł-zmiennych”?

s.71, w.3, g – Algorytm nie „charakteryzuje się następującą konstrukcją” tylko po prostu „ma następującą konstrukcję”.

s.71, w.5, g – Oczywiście p zmiennych losujemy nie „ze zbioru danych” (Dane to zbiór informacji o próbie lub populacji) tylko ze zbioru zmiennych.

s.71, w.8, g – Co to jest „głębokość reguły”? Co zostanie przekodowane na zmienną – głębokość czy reguła? Jak będzie wyglądało takie przekodowywanie?

s.71, w.10, g – „należy zakodować reguły spełniające wymienione wyżej warunki do postaci zmiennych binarnych”. Problem w tym, że „wyżej” nie ma wymienionych żadnych (!) warunków

s.71, w.12, g – Zły szyk zdania. Ma być „... a zmiennymi utworzonymi przez zaproponowany algorytm”

s.71, w.15, g – Istotne zmienne zostają włączone do zbioru zmiennych, a nie do zbioru danych. Pytanie czy oryginalne zmienne, które były wykorzystane przy tworzeniu tej nowej, pozostają w zbiorze zmiennych?

s.71, w.13, d – Zdanie „Jest to ekspercka decyzja badacza” brzmi pretensjonalnie. Trzeba było po prostu napisać, że zdecydowano, iż liczba zmiennych oryginalnych wykorzystywanych do tworzenia nowych zmiennych zawierać się będzie pomiędzy 3 a 6. Chciałbym podkreślić, że uważam ten wybór za dobry i uzasadnienie podane w tym akapicie jest słuszne.

s.71, w.3, d – Zły szyk zdania. Ma być „W przypadku badania przeprowadzonego w rozdziale 5 ...”

s.72, w.12, g – Chyba znów zły szyk zdania – wyraz „podział” jest na końcu, zamiast po wyrazie „jedynie”. Nie bardzo rozumiem czym różni się w omawianym kontekście podział od klasyfikacji.

s.72, w.13, d – Znów złe położenie wyrazu. „badaniu” ma być po „w”, a nie przed „pozwała”

s.72, w.10, d – Znów Autor nie rozróżnia cech od ich wartości

s.73, w.4, g – Dziwne „spreparowane przykłady”, a to są po prostu umowne (lub przykładowe) kategorie klientów

s.73, w.16, d – Nie ma czegoś takiego jak „logika rozkładu zmiennych ciągłych” więc nie może na niej opierać się algorytm. Dodatkowo „rozkład” jest w liczbie pojedynczej, a „zmienne” w mnogiej co sugeruje zmienną losową wielowymiarową

s.73, w.15, d – Chyba chodzi o kombinację „wybranych wariantów cech nominalnych”?

s.73, w.1, d – „grupy odniesienia” powinno być przed „odpowiadającej”

s.74, w.1, g – Oczywiście wyznaczana jest „ocena dystrybuanty” (lub dystrybuanta empiryczna), a nie dystrybuanta

s.76, w.1, g – Nie „kombinacji cech” tylko „kombinacji wartości cech”

s.77, w.10, d – Żargonowe „okno czasowe”

s.80, Podpis pod Tabelą 2 – W Słowniku Języka Polskiego PWN są wymienione następujące znaczenia słowa „aplikować: 1. «odbywać aplikację – praktykę», 2. «stosować lub podawać lek albo kosmetyk; też: zalecać go lub przepisywać», 3. «narzucać coś komuś, stosować coś wobec kogoś», 4. «naszywać aplikacje». Klient więc nie „aplikuje o kredyt” tylko „stara się” lub „wnioskuje”. To samo: s.101, w.8, g

s.80, punkt 4 – Żargonowa „flaga”, a chodzi tu o wartość 1 odpowiedniej zmiennej zero-jedynkowej

s.81, w.2, g – „wniosków” ma być po „wszystkich”

s.81, w.81, 6 – Straszne zdanie „Algorytmy zostały zaaplikowane na przedstawionych w podrozdziale 5.1 danych, na których opiera się badanie ...”. Algorytmy zostały zastosowane

do analizy danych, a nie mogły być zaaplikowane, bo nie są lekarstwami. Dane nie zostały „przedstawione” w podrozdziale 5.1 (nikt nie wymaga podawania ponad 24 tysięcy rekordów), tylko zostały tam „opisane”. Badanie nie „opiera się na danych” tylko je „wykorzystuje”.

s.81, w.13, d – Rozwiązanie standardowe nie może „opierać się” na regresji logistycznej.

s.81, w.9, d – Wypadało choćby w kilku słowach opisać miarę Lift.

s.81, w.8, d – „w sekcji 5.3.1” to język EngPol. Po polsku chodzi tu nie o „sekcję” tylko o część, punkt lub ustęp. Słownik Języka Polskiego PWN podaje następujące znaczenia słowa „sekcja”: 1. «dział urzędu, organizacji itp.», 2. «badanie zwłok w celu stwierdzenia zmian chorobowych i ustalenia przyczyny zgonu», 3. «część urządzenia mająca określony zakres działania», 4. «część plutonu», 5. «grupa żołnierzy obsługujących określony sprzęt bojowy». To samo: s.92, w.4, g; s.95, w.16, d; s.96, w.4, g; s.103, w.8, 9 i 16, d; s.104, w.5, g; s.109, w.17, g; s.114, w.3, 11 i 12, d; s.115, w.9, 10, 15 i 16, g

s.81, w.2, d – „inicjalny” to czytany „po polsku” angielski wyraz „initial” (czyli EngPol w czystej postaci), a chodzi pewnie o wyjściowy zbiór zmiennych (lub zmienne oryginalne). To samo: s.82, w.3, g; s.82, w.9, g; s.83, w.3, d; s.87, w.3, d; s.93, w.3, 4 i 7, d; s.94, w.1, 9,11 i 13, g; s.96, w.12, d

s.82, w.3, g – Zbiór jest przekształcany „w” inny zbiór, a nie „na”

s.82, w.8, g – Straszne zdanie „Zastosowano trzy wersje eksperymentu, zakładające oparcie głównych składowych na trzech transformacjach zbioru inicjalnego, adresujących problem informacyjnych braków danych”. Proszę o przetłumaczenie tego zdania na język polski. Wersji eksperymentu nie można „zastosować”, główne składowe na niczym się nie opierają, zbiór jest wyjściowy, on podlega przekształceniom (jakim?), czy mogą być braki danych inne niż „informacyjne”

s.82, w.12, g – Zwracam uwagę, że po pierwsze – percentyle to liczby, a nie przedziały liczbowe; po drugie – percentyli jest 99; więc o które „5 kolejnych” chodzi.

s.82, Tabela 3 – To nie zbiór jest imputowany, tylko braki danych. To samo: s.88, Tabela 12

s.83, w.14, g – Angielski wyraz „finally” tłumaczymy nie jako „finalnie” (bo to jest EngPol, a nie tłumaczenie) tylko jako „w końcu” „ostatecznie”, a użytym kontekście „w sumie”. To samo: s.89, w.11, g; s.96, w.14, d; s.101, w.1 i 7, g

s.88 - Pierwszy akapit jest powtórzeniem informacji podanej już uprzednio

s.93, w.9, g – „training set” to „zbiór uczący”, a „validation set” to „zbiór testowy”. „Trenujący” to może być zawodnik.

s.94, w.7, d - określenie „dla 5 górnych percentyli rozkładu prognoz jest błędne”. Ma być: „dla wartości większych od 95 percentyla w rozkładzie prognoz”. To samo: s.95, w.5 i 12, g; s.96, w.5, g; s.97, w.4, g; s.98, w.15, g; s.99, w.2 i 6, d; s.103, w.3, d; s.104, w.2, 11 i 18, g; s.106, w.4 i 17, g; s.110, w.6, g oraz Podpisy pod Tabelami 14, 16, 18, 19, 20, 22; s.117, w.10, g

s.95, w.9, d – Brak litery w „undersampling”. Ale tego nie można odmieniać po polsku!

s.104, drugi wiersz nad Tabelą 19 – Niezręczne „... na przestrzeni przeprowadzonych symulacji”

s.106, w.14, g – „backward” to regresja „wsteczna” lub „zstępująca”

s.111, w.8, d – Zły szyk zdania. Wyraz „rozwiązania” powinien być po „stabilności, a nie na końcu zdania.

s.114 – Sądzę, że niepotrzebnie powtórzono hipotezy, niestety łącznie z ich błędami logicznymi. „Zdarzeń rzadkich nie można prognozować za pomocą nowego algorytmu konstrukcji zmiennych”, bo on służy (jak sama nazwa wskazuje) do konstrukcji zmiennych, a nie do prognozowania.

s.119, w.4, d – „Joint” powinno być z wielkiej litery

s.120, w.9, g – „internet” powinno być z małej litery (skoro kino, teatr, radio, telewizja są z małej)

oj. Lelonek