

Пустовіт Дмитро Тарасович

студент

Національного технічного університету України

«Київський політехнічний інститут імені Ігоря Сікорського»

Пустовит Дмитрий Тарасович

студент

Национального технического университета Украины

«Киевский политехнический институт имени Игоря Сикорского»

Pustovit Dmytro Tarasovych

Student of the

National Technical University of Ukraine

«Igor Sikorsky Kyiv Polytechnic Institute»

МАРКІВСЬКИЙ ПІДХІД В ЗАДАЧІ ПЕРЕДБАЧЕННЯ КРИЗОВИХ ЯВИЩ РИНКУ ПРАЦІ В УКРАЇНІ

МАРКОВСКИЙ ПОДХОД В ЗАДАЧЕ ПРЕДВИДЕНИЯ КРИЗИСНЫХ ЯВЛЕНИЙ НА РЫНКЕ ТРУДА УКРАИНЫ

PREDICTING LABOR MARKET DOWNTURNS IN UKRAINE USING MARKOV APPROACH

Анотація. Досліджено можливість вирішення проблеми прогнозування кризових явищ шляхом побудови приховано-го марківського ланцюга. Створено систему раннього сповіщення для криз ринку праці в Україні.

Ключові слова: прихований ланцюг маркова, система раннього сповіщення, рівень безробіття, криза.

Аннотация. Исследована возможность решения проблемы прогнозирования кризисных явлений путем построения скрытой марковской цепи. Создана система раннего оповещения для кризисов рынка труда в Украине.

Ключевые слова: скрытая цепь Маркова, система раннего оповещения, уровень безработицы, кризис.

Summary. The possibility of solving the crisis prediction problem based on hidden Markov models was investigated. An early warning system for labor market crises in Ukraine was created.

Key words: hidden markov models, early warning system, unemployment rate, crisis.

Постановка проблеми. Історично людина в процесі своєї діяльності прагне уникати хаосу, непередбачуваних ситуацій та нестабільності. Сьогодні як ніколи світ потребує ефективного засобу виявлення та боротьби з кризовими явищами, в першу чергу в таких сферах як економіка, міжнародні відносини, соціальна сфера, екологія. Втім, кризу можна розглядати і як переломний момент у розвитку системи, який дає простір для нового витка економічних змін. З такої точки зору особливо корисним буде існування системи раннього сповіщення про кризові явища, яке дозволило б максимально контролювати кризові процеси та відвертати

катастрофічні наслідки можливих криз ринку праці для економіки та суспільства. Цій проблемі було присвячено безліч різноманітних праць, проте більшість моделей виявляються неефективними в умовах ризиків, невизначеності інформації чи стрибкоподібних змін.

Мета статті. В цій роботі пропонується зосередити увагу на альтернативному — Марківському підході, тобто на побудові моделей, для яких діє марківське припущення про незалежність станів. Цим шляхом пішло досить багато науковців, досліджуючи проблеми різного характеру в різних галузях, зокрема розпізнавання мови та штучного перекладу.

Аналіз останніх досліджень і публікацій. Однією з найпотужніших марківських моделей, використаній і в даній роботі є приховані марківські ланцюги. Вони були успішно застосовані щодо військових конфліктів, зокрема точність короткострокових прогнозів для війни в Югославії (у період з січня 1991 по січень 1999) досягала 92% [5, с. 29]. Також ця модель отримала застосування в проблемах міграції [3, с. 25–30], аналізі ризиків при формуванні інвестиційного портфелю [1, с. 5–10], аналізі фондового ринку [6, с. 32–45], навіть для передбачення результатів виборів [2, с. 26–30].

Виклад основного матеріалу. Основою запропонованої моделі стане прихований марківський ланцюг з п'ятьма прихованими станами: велике падіння, мале падіння, ринок без змін, мале зростання та велике зростання (кризове явище). В даній роботі для навчання було обрано місячні дані показників безробіття в Україні за 2010–2017 роки. Додамо ще 5 станів, які будуть спостережуваними, і використовуватимуться для класифікації початкових даних.

Формально таку модель представляють у вигляді структури: $\{S, V, \Pi, A, B\}$, де:

- 1) $S = \{1, \dots, 5\} = \{s_1, \dots, s_5\}$ — це множина станів.
- 2) $V = \{v_1, \dots, v_M\}$ — це вихідний алфавіт. В даному випадку M — це кількість можливих значень які ми можемо спостерігати: $V = \{1, 2, 1, 3, \dots, 2, 1, 2, 2\}$.
- 3) $\Pi = \{\pi_i\}_{i \in S}$ — початковий розподіл, тобто $\pi_i = \mathcal{P}\{s_1 = i\}$.
- 4) $A = \{A_j^i\}_{i, j \in S}$ — матриця ймовірнісних переходів (матриця перехідних імовірностей), тобто:
 $A_j^i = \mathcal{P}\{s_{t+1} = i | s_t = j\}$, $1 \leq i, j \leq 5$.
- 5) $B_{x_t}^j = \{b_j(x_t)\}$ — розподіл ймовірності спостережень, тобто:

$$B_{x_t}^j = \mathcal{P}\{x_t = k, s_t = j\}, k \in V.$$

Навчання моделі здійснюватиметься за варіацією EM-алгоритму для прихованих Марківських ланцюгів — алгоритму Баума-Велша [4, с. 11–17]:

Спершу слід ввести 2 додаткові змінні:

$$\xi(t, i, j) = \mathcal{P}\{s_t = i, s_{t+1} = j | \bar{x}, \mu\} = \frac{\mathcal{P}\{s_t = i, s_{t+1} = j, \bar{x} | \mu\}}{\mathcal{P}\{\bar{x} | \mu\}} = \frac{\alpha_i(t) * A_j^i * B_{x_t}^i * \beta_j(t+1)}{\sum_{m=1}^N \alpha_m(t) \beta_m(t)} = \frac{\alpha_i(t) * A_j^i * B_{x_t}^i * \beta_j(t+1)}{\sum_{m=1}^N \sum_{n=1}^N \alpha_m(t) * A_j^i * B_{x_t}^i * \beta_n(t)}$$

$$\gamma_i(t) = \mathcal{P}\{s_t = i | \bar{x}, \mu\} = \sum_{j=1}^N \mathcal{P}\{s_t = i, s_{t+1} = j | \bar{x}, \mu\} = \sum_{j=1}^N \xi(t, i, j)$$

Тепер формули пошуку оптимальних параметрів моделі в нових позначеннях виглядають так:

$$\bar{\pi} = \mathcal{P}\{s_0 = i | \bar{x}, \mu\} = \gamma_i(0)$$

$$\tilde{A}_j^i = \frac{\sum_{t=1}^T \mathcal{P}\{s_t = i, s_{t+1} = j | \bar{x}, \mu\}}{\sum_{t=1}^T \mathcal{P}\{s_t = i | \bar{x}, \mu\}} = \frac{\sum_{t=1}^T \xi(t, i, j)}{\sum_{t=1}^T \gamma_i(t)}$$

$$= \frac{\sum_{t=1}^T \alpha_i(t) * A_j^i * B_{x_{t+1}}^i * \beta_j(t+1)}{\sum_{t=1}^T \alpha_i(t) * \beta_i(t)}$$

$$\tilde{B}_k^i = \frac{\sum_{t=1}^T 1\{(x_t \in k)\} * \mathcal{P}\{s_t = i | \bar{x}, \mu\}}{\sum_{t=1}^T \mathcal{P}\{s_t = i | \bar{x}, \mu\}} = \frac{\sum_{t=1}^T 1\{(x_t \in k)\} \gamma_i(t)}{\sum_{t=1}^T \gamma_i(t)} \quad (2.10)$$

$$= \frac{\sum_{t=1}^T 1\{(x_t \in k)\} * \alpha_i(t) * \beta_i(t)}{\sum_{t=1}^T \alpha_i(t) * \beta_i(t)}$$

$\alpha_i(t)$ та $\beta_i(t)$ тут являють собою наступні ймовірності:

$$\alpha_i(t) = \mathcal{P}\{x_1, x_2, \dots, x_T, z_t = s_i, A, B\}$$

$$\beta_i(t) = \mathcal{P}\{x_T, \dots, x_1, z_t = s_i, A, B\}.$$

Їх можна знайти безпосередньо, або ж використати алгоритми динамічного програмування. Наприклад пошук $\alpha_i(t)$ здійснюється з допомогою алгоритму Прямого Проходу (англ. Forward Procedure):

Ініціалізація:

$$\alpha_i(0) = \pi_i, i = 1..|S|.$$

Рекурсія:

$$\alpha_j(t) = \sum_{i=1}^{|S|} \alpha_i(t-1) * A_j^i * B_{x_t}^i, j = 1..|S|, t = 1..T.$$

Повністю аналогічну процедуру пошуку $\beta_i(t)$ називають Зворотнім Проходом (Backward Procedure).

Для прогнозів переходу в кризовий стан можна було б використати класичний алгоритм Вітербі, але він недостатньо потужний для вирішення цієї вузької задачі. У зв'язку з цим прогнозування здійснюватиметься за лобовим методом:

$$\mathcal{P}\{в\ найближчі\ n\ кроків\ криза\} = \sum_{i=1}^n \mathcal{P}\{криза\ настане\ на\ i\text{-му\ кроці}\} = \sum_{i=1}^n \sum_{s_j \in S, s_2 \in S, \dots, x_{i-1} \in S, z_i = s_i}$$

Це не буде настільки швидкий алгоритм, як алгоритм Вітербі, але для короткострокових прогнозів даний підхід дає кращі результати.

Відповідно, застосувавши такий підхід до даних ринку праці України 2010–2017 років матимемо таку систему раннього сповіщення про кризові явища, зображену в часовому розрізі на Рисунку 1:

Для кожної з цих п'яти областей було обчислено ймовірності переходу в кризовий стан в короткостроковий період (Таблиця 1). Кризою при цьому вважаємо таку ситуацію, де значення досліджуваного індексу перевищує середнє значення більше ніж на одне стандартне відхилення.

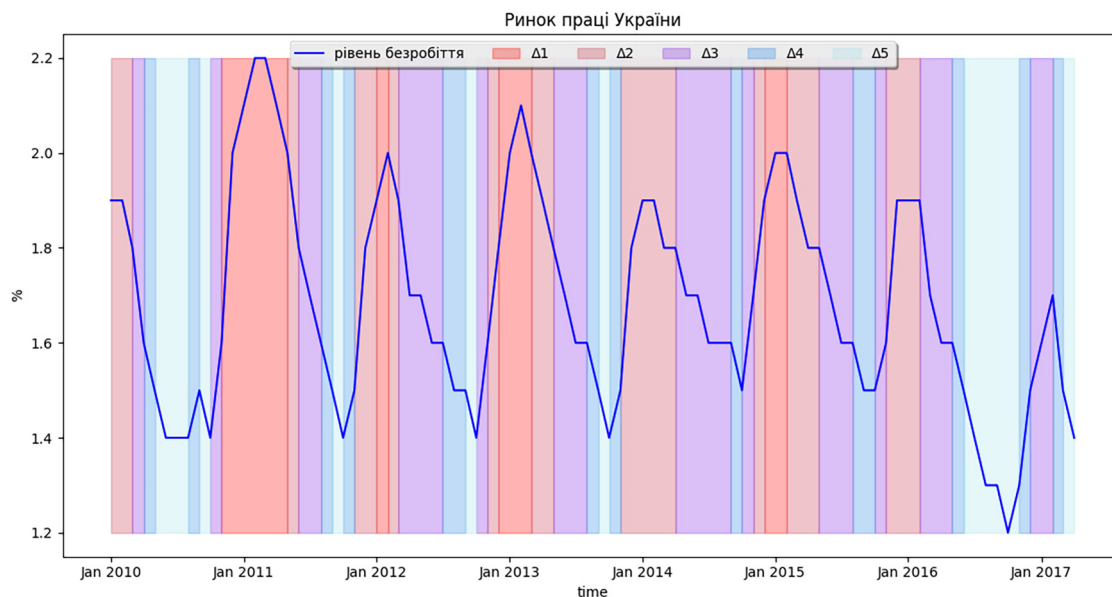


Рисунок 1. Система раннього сповіщення (Розробка автора)

Таблиця 1

**Ймовірність короткострокової кризи
(розробка автора)**

Область	Ймовірність настання Кризи
Δ_1	0.7364154270871943
Δ_2	0.2660487971938277
Δ_3	0.10102797022557222
Δ_4	0.08134463978183368
Δ_5	0.0255177875929092

Висновки. Отже, як було показано задача прогнозування кризових явищ може бути успішно вирішена з допомогою марківського підходу. Побудований прихований марківський ланцюг дозволив створити систему раннього сповіщення про кризові явища ринку праці України на основі показників безробіття 2010–2017 років. Відповідно станом на червень 2017 року ймовірність настання кризи ринку праці України в короткостроковому періоді становить лише 2,5%. Такий результат лише підтверджує загальну тенденцію до зниження рівня безробіття за останні кілька років.

Література

1. Hidden Markov Model of Portfolio Credit Risk / [S. Ahuja,, S. Bunsupha,, K. Tan and el.]. — Stanford, 2011. — 17 p.
2. Wagner. C. S. Presidential Election Forecasts: Through the Lense of Linear Algebra [Електронний ресурс] / Cassia S. Wagner // Electronic Thesis and Dissertation Repository. — 2012. — Режим доступу до ресурсу: <http://docplayer.net/26233199-U-s-presidential-election-forecasts-through-the-lense-of-linear-algebra-cassia-s-wagner.html>.
3. Constant A. The Dynamics of Repeat Migration: A Markov Chain Analysis / A. Constant, K. Zimmermann. — Bonn: CEPR, 2004. — 39 p.
4. McCallum A. Hidden Markov models Baum Welch algorithm [Електронний ресурс] / Andrew McCallum // College of Information and Computer Sciences. — 2004. — Режим доступу до ресурсу: <https://people.cs.umass.edu/~mccallum/courses/inlp2004a/lect10-hmm2.pdf>.
5. Schrodt P. A. Pattern Recognition of International Crises using Hidden Markov Models [Електронний ресурс] / Philip Schrodt Schrodt // Parus Analytical Systems. — 1997. — Режим доступу до ресурсу: <http://eventdata.parusanalytics.com/papers.dir/schro97b.pdf>
6. Zhang Y. J. Prediction of financial time series with hidden Markov models [Електронний ресурс] / Ying jian Zhang // Simon Fraser University. — 2001. — Режим доступу до ресурсу: https://www.cs.sfu.ca/~anoop/students/rzhang/rzhang_msc_thesis.pdf.