

Кінда В. В., аспірант (Рівненський державний гуманітарний університет, м. Рівне)

АНАЛІЗ ЯКОСТІ ПРОГНОЗУ НА ОСНОВІ ЗАСТОСУВАННЯ НЕЧІТКОЇ ЛОГІКИ

Реалізовано метод, що базується на застосуванні нечіткої логіки в середовищі розробки високого рівня Python у вигляді програмного продукту. Наведено приклад визначення універсальної множини для часових рядів, побудови функцій належності на інтервалах універсальної множин та способи її задання. Проведено порівняльний аналіз короткострокового точкового прогнозу методом нечіткої логіки з пірамідальним методом, в основі якого лежить аналіз розділених різниць. Проведено апробацію роботи програмного продукту та проаналізовано одержані результати.

Ключові слова: часовий ряд; нечітка множина; функція належності; екстраполяція; прогноз.

Актуальність теми. Із дедалі швидшим темпом розвитку та популяризацією нового ринку криптовалют все більшої актуальності набуває аналіз часових рядів. Водночас зростає кількість юридичних та фізичних осіб, основною задачею яких є торгівля з метою отримання прибутку на різного роду платформах, біржах. Таким чином, розробка та застосування систем короткострокового прогнозу економічних часових рядів у процесі прийняття рішення щодо здійснення фінансових операцій є нині актуальною.

Аналіз досліджень. За останні роки дослідження часових рядів для точкового та інтервального прогнозів з використанням методів інтелектуального аналізу даних розвинулось в окрему галузь. Велику кількість займають методи на основі нейронних мереж, генетичних алгоритмів [1]. Проте згадані моделі складні у використанні та ресурсоємні, на відміну від моделей пірамідального методу. Простота та швидкість одержання результату спонукає до необхідності в подальшому дослідженні пірамідального методу екстраполяцій для точкового прогнозу та аналізу методу на основі нечіткої логіки.

Метою досліджень є реалізація методу пірамідальної екстраполяції моделей, базованих на застосуванні нечіткої логіки для оціню-

вання ефективності короткострокового прогнозу на базі обраного критерію у вигляді програмного продукту.

Викладення основного матеріалу. Якщо неможливе використання методів машинного навчання (базованих на нечітких системах), можна використовувати системи на базі нечіткої логіки. До їхніх ключових особливостей належать [2]:

- керованість;
- простота реалізації;
- масштабування.

Проблема дослідження будь-якого ряду полягає у виокремленні максимальної кількості інформації від досліджуваного процесу. У найпростішому випадку, щоб зробити прогноз, можна побудувати криву по наявних точках в певному класі кривих функцій $F(t)$. Проте зі збільшенням кількості спостережень такий підхід може призвести до ускладнення моделі та, як наслідок, – до похибок в прогнозі.

Деякою мірою побудова моделі прогнозу – це задача відокремлення сигналу від шуму, загальної поведінки від локальних особливостей та випадкових помилок, детермінованої компоненти від випадкової компоненти процесу.

Якщо маємо часовий ряд Y , його окремі значення $y(t) \in Y$ та оцінку для Y , функція $\hat{y}(t+1) = f(y(t))$, то метою f є мінімізація $\varepsilon(t)$, де $\varepsilon(t) = y(t) - \hat{y}(t)$ при $t \rightarrow \infty$, $\lim \varepsilon(t) = 0$.

Нижче наведемо параметри, від яких найбільше залежить точність моделі FTS.

1. Кількість інтервалів розбиття універсальної множини. Є найбільш впливовим параметром точності моделі, який визначається здебільшого емпіричним шляхом із врахуванням проблеми перенавчання (рис. 1).

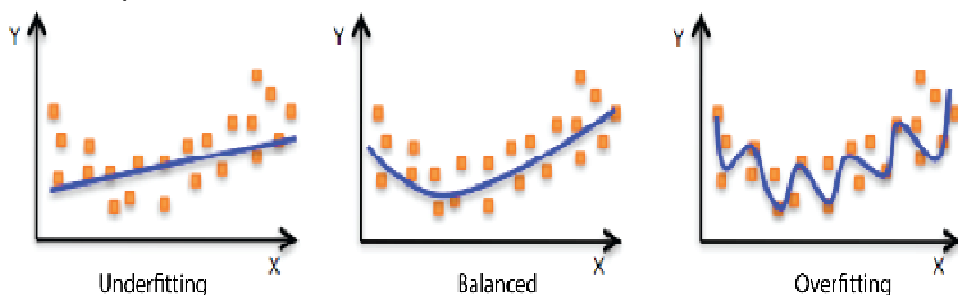


Рис. 1. Приклад побудови прогнозної функції в залежності від якості навчання

2. Тип функції належності. Це параметр, який мало впливає на точність моделі, але залежно від випадку може бути вагомою причиною для використання функції Гауса чи трапеції замість трикутної функції належності, яка є типовою.

Одним із обґрунтувань може бути кількість параметрів (гаусівський використовує 2, трикутний 3 і трапецієподібний 4), розбірливість моделі або навіть інші питання, пов'язані з природою процесу та даних.

3. Тип моделі.

Weighted – вагові коефіцієнти підвищують точність, балансування яких у правилах моделі [5] дає більш якісний прогноз за обраними критеріями оптимізації.

При апробації роботи на реальних даних було використано моделі HOFTS (без вагів), WHOFTS (з консеквентними вагами) і PWFTS (з прецедентними та консеквентними вагами).

У більшості випадків ми маємо лише часові ряди з однією змінною – ендогенною змінною. В інших випадках цій змінній допомагає інша інформація (екзогенні змінні), якою ми можемо скористатися.

Якщо часовий ряд містить багатовимірні дані, то перше, що потрібно знати, чи існує кореляція між змінними. Коефіцієнт кореляції вказує на просту лінійну залежність, тому він не повинен бути унікальним інструментом, який можна використовувати, кросентропія є хорошою альтернативою.

Результати роботи. Нехай дано часовий ряд Y , його окремі значення $y(t) \in Y$. Для кожного елемента ряду можна згенерувати наступні змінні – швидкість та прискорення (швидкість зміни швидкості). Оскільки значення ряду можуть мати велике розсіювання, тому визначення швидкості необхідно проводити не в абсолютних величинах, а у відсоткових.

Тоді для кожного спостереження, починаючи із другого можна побудувати інтервал, в центрі якого значення спостереження та границі відповідають умовам $[y_i - v_i, y_i + v_i]$. Саме ці інтервали можна використати в якості функцій належності для прогнозу. Якщо кількість спостережень більша ніж необхідна кількість інтервалів, можна використовувати інтелектуальне групування задачі кластеризації.

Для проведення практичних експериментів прогнозування взяли дані із джерела <http://sonda.ccst.inpe.br/>, графік наведено на рис. 2.

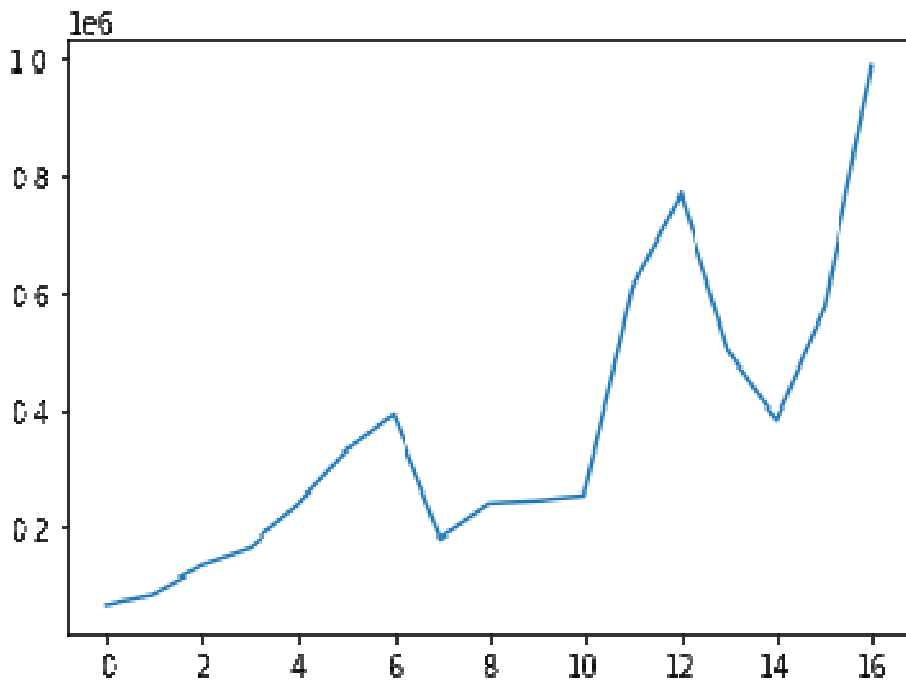


Рис. 2. Графічна візуалізація досліджуваного набору даних

Отримані результати наведемо у зведеній таблиці.

Таблиця

Результати відносних похибок прогнозу шуканої функції

	Реальні значення	Прогноз	Різниця	Відносна похибка, %
Without weights (HOFTS)	986778	1123887.3	-137109.3	13.9
With weights in consequent of the rules (WHOFTS)	986778	922345.1	64432.9	6.5
With weights in the consequent and the precedent of the rules (PWFTS)	986778	948194.2	38583.8	3.9
Piramidal	986778	924123.4	62654.6	6.3

Використання предиктивних моделей базованих на нечіткій логіці доцільно використовувати для стаціонарних часових рядів. Інакше маємо високу похибку прогнозів згідно з обраними критеріями оптимізації. Для роботи із такими рядами можна використовувати операції зведення до стаціонарності використовуючи тести перевірки її наявності. В результаті проведення короткострокового прогнозу для запропонованого часового ряду отримали наступні результати таблиці. Оптимальні параметри підбирались на сітці $[0,1]$ із кроком дискретизації 0,05. Як видно із таблиці модель PWFTS на заданому наборі даних в порівнянні із WHOFTS та HOWFTS. Це пояснюється методом оптимізації та пошуку параметрів моделі. Пірамідальний метод екстраполяції дає точність співрозмірну із результатом роботи моделі WHOFTS.

Висновки та перспективи подальших досліджень. Побудовано моделі короткострокового прогнозу із застосуванням нечіткої логіки. Моделі на базі нечіткого виводу правил Сугено без використання вагів, із консеквентними вагами та з прецедентно-консеквентними вагами. Розроблено програмний продукт у середовищі розробки високого рівня Python. Отримані результати підтверджують необхідність більш детального вивчення та модифікації даних методів, що і є напрямком наших подальших досліджень.

1. V. R. Uslu, E. Bas, U. Yolcu and E. Egrioglu. A fuzzy time series approach based on weights determined by the number of recurrences of fuzzy relations. *Swarm Evol. Comput.* Apr. 2014. Vol. 15. 2. L. A. Zadeh. Fuzzy Sets, Fuzzy Logic and Fuzzy Systems: Selected Papers by Lotfi A Zadeh. Singapore : World Scientific, 1996. 3. Q. Song and B. S. Chissom. Forecasting enrollments with fuzzy time series. *Fuzzy Sets Syst.* 1993. Vol. 54, No. 1. Part I. 4. K. Huarng. Effective lengths of intervals to improve forecasting in fuzzy time series. *Fuzzy Sets Syst.* Nov. 2001. Vol. 123, No. 3. 5. Зайченко Ю. П. Исследование операций. 2-е изд. Киев : Изд-во «Вища школа», 1979.

REFERENCES:

1. V. R. Uslu, E. Bas, U. Yolcu and E. Egrioglu. A fuzzy time series approach based on weights determined by the number of recurrences of fuzzy relations. *Swarm Evol. Comput.* Apr. 2014. Vol. 15. 2. L. A. Zadeh. Fuzzy Sets Fuzzy Logic and Fuzzy Systems: Selected Papers by Lotfi A Zadeh. Singapore : World Scientific, 1996. 3. Q. Song and B. S. Chissom. Forecasting enrollments with fuzzy time series. *Fuzzy Sets Syst.* 1993. Vol. 54, No. 1. Part I. 4. K. Huarng.

Effective lengths of intervals to improve forecasting in fuzzy time series. *Fuzzy Sets Syst. Nov.* 2001. Vol. 123, No. 3. 5. Zaichenko Yu. P. *Issledovanie operatsii*. 2-e izd. Kiev : Izd-vo «Vyshcha shkola», 1979.

Kinda V. V., Post-graduate Student (Rivne State University for the Humanities, Rivne)

ANALYSIS OF FORECAST QUALITY BASED ON THE APPLICATION OF FUZZY LOGIC

Simplicity, speed of obtaining the result, the possibility of scaling prompts the need for further research of the pyramid method of extrapolation and analysis of methods based on fuzzy logic for short-term point forecasting of financial time series.

The purpose of the research is to compare the pyramidal extrapolation method, models based on the application of fuzzy logic for evaluating the effectiveness of short-term forecasting based on a selected criterion in the form of a software product. The main criterion for checking the adequacy of the forecast is the relative error.

Implemented methods based on the application of fuzzy logic in a high-level Python development environment in the form of a software product. An example of defining a universal set for time series, construction of membership functions on intervals of a universal set, and ways of specifying it is given. A comparative analysis of the short-term point forecast by the method of fuzzy logic with the pyramidal method, which is based on the analysis of separated differences, was carried out. The software product was tested on real data and the results were analyzed.

***Keywords:* time series; fuzzy set; membership function; extrapolation; forecast.**
