

Предсказване на замърсявания в природата на базата на невронни мрежи

Христо Вълчанов, Венцислав Николов

Резюме: В статията е разгледан метод за анализ и едномерно предсказване с невронна мрежа на временни редове, представлящи измервания на показатели на природни замърсявания. Представени са някои от особеностите на реализирана система за предсказване и предимствата на метода спрямо някои от най-често използваните алтернативни методи.

Nature Analyzing and Prediction Based on Neural Networks

Hristo Valchanov, Ventsislav Nikolov

Abstract: In this paper an approach for analysis and univariate time series prediction of nature pollutions by a neural network is proposed. Some specific features of the realized system are presented, and advantages of the approach compared to some of the most commonly used alternative approaches are presented.

1. Въведение

В природата ефектът от едни или други явления и човешки дейности обикновено не се проявяват веднага, а в продължение на определен период от време. Това дава възможност за използване на определени статистически методи за анализиране и предсказване на състоянието на определени процеси свързани със замърсяването на природата. От друга страна състоянието на определен показател почти винаги се измерва през определен постоянен интервал от време. Така поредицата от дискретни измервания на даден показател образуват временен ред и могат да се използват някои от множеството известни методи за предсказване на временни редове. Тук се разглежда т.нар. едномерно предсказване, т.е. предсказване при което за да се получи определена бъдеща стойност се анализират само наличните исторически стойности за този ред без да се отчита влиянието на външни фактори. Причината е, че в природата влияят много фактори и е трудно да се отчете с точност всеки от тях както и да

се комбинират според тяхната тежест. В настоящата статия е представен подход за приложение на невронна мрежа за предсказване на екологични замърсявания. Обърнато е внимание на предимствата на реализацията като нелинеен метод пред класическите статистически подходи.

2. Описание на проблема

Някои от по-известните и често използвани методи за анализ и предсказване на временни редове са: авторегресионни методи и техни комбинации с пълзящи средни – AR, ARMA, ARIMA и т.н, метод на Holt-Winter, методи на декомпозиция, експоненциално изглаждане и др.[1,3] По-долу са представени накратко авторегресионните модели, които като идея най-много наподобяват принципа на едномерно предсказване на временен ред с невронна мрежа. Тези модели са полезни когато състоянието на определен показател може да се представи като линейна комбинация от краен брой негови предшестваци стойности и случайни грешки.

2.1. Авторегресионен – AR (AutoRegressive)

$$y(t) = \varphi_0 + \varphi_1 y(t-1) + \varphi_2 y(t-2) + \dots + \varphi_p y(t-p)$$

където $\varphi_0, \varphi_1, \dots, \varphi_p$ са параметрите на модела.

Моделът се определя от броя на елементите p . Използването на този метод изисква временните редове в периодите на изследване да са стационарни.

2.2. Пълзящи средни – MA (Moving Average)

$$y(t) = a_t - \theta_1 a(t-1) - \dots - \theta_q a(t-q)$$

където $\theta_1, \theta_2, \dots, \theta_q$ са параметрите на модела;

$a(t-1), a(t-2), \dots, a(t-q)$ – бял шум.

Тези модели имат по-ограничено приложение и могат успешно да се използват само ако редът от случайни колебания се подчинява на някакви закономерни последователности.

2.3. Авторегресивни и пълзящи средни – комбинация от горните два метода – ARMA (AutoRegressive Moving Average)

$$y(t) = \varphi_0 + \varphi_1 y(t-1) + \varphi_2 y(t-2) + \dots + \varphi_p y(t-p) + a_t - \theta_1 a(t-1) - \dots - \theta_q a(t-q)$$

където $\varphi_0, \varphi_1, \dots, \varphi_p$ са параметрите на AR модела, а $\theta_1, \theta_2, \dots, \theta_q$ са параметрите на MA модела.

Тези процеси притежават характеристиките на предходните два.

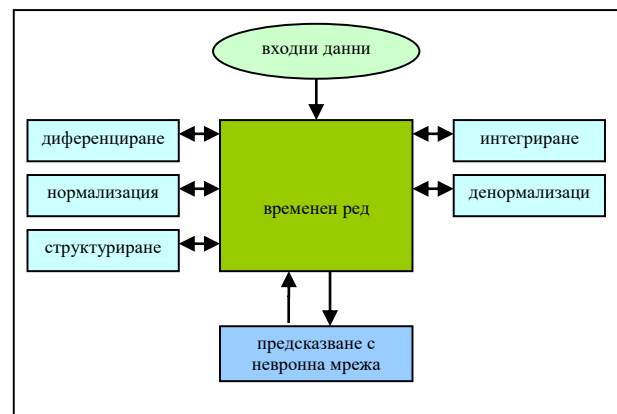
Основните недостатъци на тези методи се състоят в това че не винаги са подходящи за предсказване на всеки временен ред. Освен това в редица случаи се налага значителна предварителна обработка на реда. Точността на резултатите може да се контролира предимно, чрез промяна на броя елементи от които се приема че

зависи следващата стойност, т.е. модела на метода.

Необходимостта от използване на по-мощен подход се поражда от факта, че природните процеси обикновено не са линейни и пропорционални, а асиметрични и нелинейни. Необходимо е средство което да дава по-голяма свобода и повече параметри при извършване на предсказването. Поради това е анализиран подходът с невронни мрежи, тъй като те могат да апроксимират поведението на произволен временен ред по интуитивен и както показват резултатите в повечето случаи достатъчно ефективен начин.

3. Предлаган подход

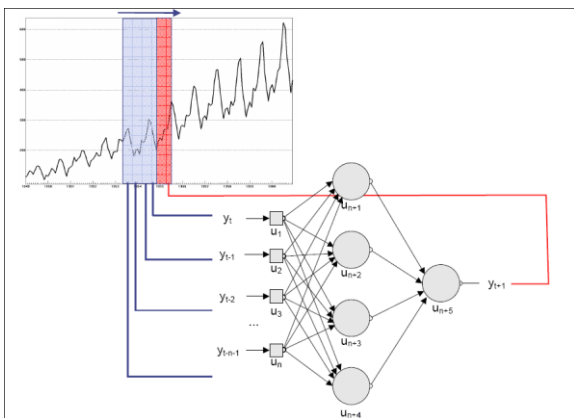
При невронните мрежи смисълът на изчислените коефициенти на теглата на връзките между отделните изчислителни елементи не се интерпретира, тъй като тези коефициенти не носят полезна информация за хората. Невронната мрежа може да се разглежда като черна кутия на която се подават определени примери и впоследствие генерира резултати на базата на тези примери (фиг.1). В този смисъл колкото повече са представените примери толкова е по-голяма вероятността за получаване на по-точни резултати. При повече примери обаче е необходимо и да се избере достатъчно мощна архитектура която да е в състояние да заучи тези данни.



Фиг.1 Обработка на информацията в системата

Едно от най-големите предимства на невронната мрежа пред класическите статистически подходи е, че тя може да открива нелинейни зависимости в наличните данни. Така тя е в състояние да действа като универсален апроксиматор.

При невронните мрежи както и при авторегресионните методи е необходимо да се определи броят на елементите попадащи в “плъзгащия се прозорец” (фиг.2), т.е. определяне на модела [2]. За целта в разработката се използват подобни подходи както при авторегресионните методи – изчисляване на корелационна и частна автокорелационна функция и намиране на областта при която тя се установява близо до нулевата стойност.



Фиг.2 Обхождане на временен ред с “плъзгащ се прозорец” и обучение на невронната мрежа

Характерно за повечето архитектури на невронните мрежи е нормализацията на данните преди обучението. В реализацията тя се извършва по следния начин:

$$y' = \min(y') + (y - \min(y)) \frac{p}{\max(y) - \min(y)}$$

където:

$\min(y')$ – долна граница на интервала на нормализация;

p – интервал на нормализация (за двоичен сигмоид – $[0, 1]$, за биполярен сигмоид и хиперболичен тангент – $[-1, 1]$);

y – елемент на временния ред.

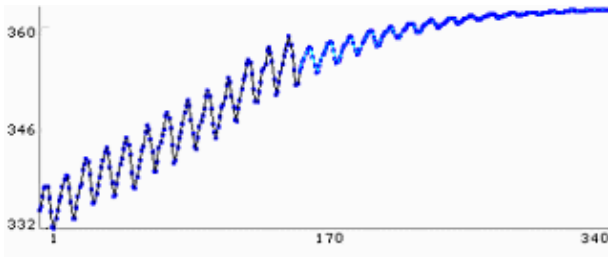
В повечето случаи се препоръчва нормализацията да се извършва не в строго определените граници $(0,1)$ или $(-1, 1)$, а в някакъв по-малък интервал, например $(0,1 - 0,9)$.

Денормализацията се извършва след анализа и прогнозирането по следния начин (означенията са същите както по-горе):

$$y = \min(y) + (y' - \min(y')) \frac{\max(y) - \min(y)}{p}$$

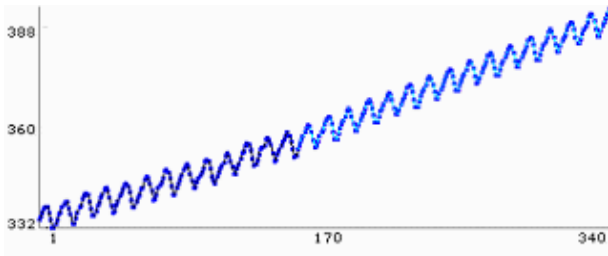
В случай че е необходимо да се денормализира предсказана стойност това се извършва на базата на минималната ($\min(y)$) и максимална ($\max(y)$) стойности на наличните исторически данни.

Въпреки че нормализацията се извършва в по-малки граници в сравнение с границите на активационната функция, при силно изразен тренд възникват някои проблеми. Тъй като дефиниционната област на активационната функция използвана при невронните мрежи е в интервала $[-1, 1]$, то изходните стойности на мрежата не могат да преминат тези граници. Така предсказаните стойности няма да могат да преминат границите породени от наличните исторически данни спрямо които е направена нормализацията. В резултат на това се получава неправилно предсказване. На фиг.3. това е демонстрирано за предсказване на временен ред от месечните стойности на CO_2 в Мауна Лоа, Хавай за периода от януари 1959г. до декември 1990г. (липсващите измервания за временния ред са попълнени чрез линейна интерполация) измерени в “части на милион” (ppm).



Фиг.3 Предсказване без елиминирание на тренд

Преди да се приложи алгоритъма на невронната мрежа се използва подход характерен за авторегресионните методи – елиминирание на съществуващия тренд на временния ред. За тази цел в разработката е използван един от най-често използваните методи – диференциране преди анализа и интегриране след получаване на резултатите. Резултатът е показан на фиг.4.



Фиг.4 Предсказване с елиминирание на тренд

Диференцираният ред е с един елемент по-малък от първоначалния при еднократно диференциране и с два при двукратно диференциране. Всеки елемент на диференцирания ред се ичислява като от съответния елемент на първоначалния ред се изважда предходния:

$$\nabla y(t) = y(t) - y(t-1)$$

където ∇ е оператор за изчисляване на разлика [4], $y(t)$ и $y(t-1)$ са стойностите от временния ред в моменти t и $t-1$ съответно.

Понякога и диференцираният ред е нестационарен при което се налага

второ диференциране по следния начин:

$$\nabla^2 y(t) = (y(t) - y(t-1)) - (y(t-1) - y(t-2))$$

По същия начин протича обработката и при ARIMA методите откъдето произлиза и наименованието им – авторегресионни интегрирани пълзящи средни (**Auto Regressive Integrated Moving Average**), поради необходимостта от интегриране преди получаване на резултатите.

След анализа и предсказването е необходимо възстановяване на тренда (интегриране). Тази операция се извършва за целия ред съдържащ оригинални и предсказани стойности. Първата стойност на възстановения ред е равна на оригиналната от оригиналния ред. Всеки следващ елемент от реда се намира по следния начин:

- първи интеграл:
 $y(i) = \nabla y(i-1) + \nabla y(i)$
- втори интеграл:
 $y(i) = \nabla y(i) - y(i-2) + 2y(i-1)$

За разлика от алтернативните методи при метода с невронна мрежа не е необходимо премахването на сезонната компонента.

Изходът на мрежата се изчислява по следния начин:

$$o(k) = f\left(\sum_{j=1}^B w(k, j) f\left(\sum_{i=1}^A w(j, i) y(i)\right)\right)$$

където:

$o(k)$ – изход на k -ти изчислителен елемент в изходния слой;

f – активационна функция;

A – брой елементи във входния слой на мрежата;

B – брой елементи в скрития слой на мрежата;

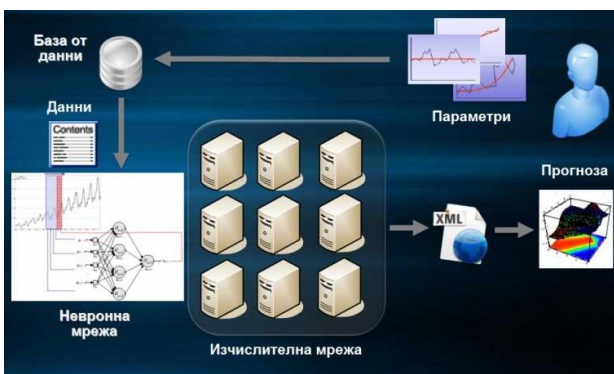
w – тегло на връзка между съответните изчислителни елементи;

$y(i)$ – елемент от временния ред.

На базата на разликата между така изчисления изход и правилния изход (най-десния елемент на “плъзгащия прозорец” на фиг.2) се настройват теглата на мрежата според правилото за обучение с обратно разпространение на грешката. След края на обучението се извършва валидация за избран интервал от наличните данни. При добри резултати обучението се приема, а при незадоволителни се отхвърля, извършват се промени на някои от настройките на мрежата и обучението се повтаря.

4. Реализация

За реализиране на предложеният подход е разработена система за предсказване. Потокът на информацията в системата е показан на фиг.5. Системата се състои от два основни модула: модул за конфигуриране на интелигентна система и модул за въвеждане на данни, анализ и предсказване. В двата модула потокът от информация е един и същ, с тази разлика, че при конфигурационния модул изчислителната мрежа извършва множество итеративни изчисления (епохи на обучение на невронната мрежа). При модула за предсказване, изчислителната мрежа извършва една итерация за едно предсказване.



Фиг.5 Поток от информация в системата за предсказване

Основната цел на модула за конфигурация е на базата на невронна мрежа да се откриват и моделират причинно-следствени връзки в природата. Модулът включва следните основни функции:

- Предварително графично изобразяване на временния ред и определяне на неговите характеристики;
- Елиминиране на тренда на реда;
- Идентификация на модела – определяне на броя на входните елементи на невронната мрежа;
- Структуриране на временния ред – подготовка за обхождане с “плъзгащ прозорец”;
- Определяне на архитектурата и параметрите на невронната мрежа – използване на определения модел, извършване на крос валидация, определяне на активационната функция, обучаваща скорост, момент и т.н.;
- Обучение на невронната мрежа;
- Валидация;
- Съхранение на резултатите и/или невронната мрежа в XML файл.

Модулът за предсказване използва създадената от конфигурационния модул невронна мрежа за генериране на предсказването, тяхното изобразяване, систематизиране и съхранение. Събраната информация за определени замърсявания се въвеждат през Интернет в централизирана база от данни. На базата на зададени от потребителя параметри, от сървър се извлича модел на определени екологични замърсявания, който се използва за генериране на прогнозни резултати.

5. Заключение

Много често при приложенията използващи невронни мрежи, качеството на резултатите силно зависи от входните данни. Необходим е подходящ набор от предварителни обработки с цел постигане на максимално точни резултати. В настоящата разработка част от тези настройки са автоматизирани, така че за тях се получават препоръчителни стойности, които могат да бъдат донастройвани и променяни динамично. Не винаги обаче съществува еднозначен критерий за качество на предсказването. Така например не винаги средно-квадратичната грешка е най-подходящ метод за измерване на точността при предсказване, тъй като интервала за нейното изчисление може да варира.

Като цяло в повечето случаи невронните мрежи притежават по-голяма гъвкавост и изчислителна мощ спрямо описаните статистически методи. Това е така, тъй като методът притежава по-голям брой параметри, което го прави по-гъвкав. В този смисъл невронната мрежа като математически модел предоставя значително по-богати възможности от алтернативните подходи, като в допълнение към основното множество от променливи тегла се използват още: междинен слой от изчислителни елементи (даващ възможност внасяне на нелинейност), отмествания (даващи допълнителна степен на свобода), избор на една от множество активационни функции (реализирани в разработката са: двоичен сигмоид, биполярен сигмоид и хиперболичен тангент) и др. Множеството параметри обаче изискват и тяхното прецизно настройване, което заедно с относително бавната скорост на обучение са най-големите недостатъци на разработката.

Представената система позволява лесно добавяне на актуални данни за природни процеси и събития,

притежава мобилност, възможности за генериране както на насочваща прогноза, така и на доверителни граници. Цел на бъдеща работа е намаляване на времето за обучение на невронната мрежа чрез използване на Grid системи за паралелна обработка.

6. Литература

1. James D. Hamilton, Time Series Analysis, ISBN: 0-691-04289-6, Princeton University Press, 1994.
2. Sven F. Crone, Business Forecasting with Artificial Neural Networks, IBF Tutorial – Institute of Business Forecasting, Boston, 2004.
3. Box G. E. P., Jenkins G. M. Time Series Analysis: Forecasting and Control, San Francisco, Holden-Day, 1970.
4. Peter J. Brockwell Richard A. Davis. Introduction to Time Series and Forecasting. Springer, 2002.

7. За авторите

Инж. Венцислав Николов
Старши Софтуерен Програμισ
Евмориск Системи ООД
9002 Варна
Ул. Генерал Киселов 31
Ел. поща:
vnikolov at eurorisksystems dot com

Гл. ас. д-р инж. Христо Вълчанов
катедра:
„Компютърни науки и технологии“
Технически Университет - Варна
Ел. поща: hristo at tu-varna dot bg