

Արհեստական նեյրոնային ցանցերը գների մոդելավորման գործընթացում

Թևոսյան Ա.

ԵՊՀ (Հայաստան, Երևան)

atevossyan@yahoo.com

Վճռորոշ բառեր՝ նախասնույց արհեստական նեյրոնային ցանցեր, մոդելի գերահարմարեցում, գների մոդելավորում

Искусственные нейронные сети в процессе анализа цены

Тевосян А.

ЕГУ (Армения, Ереван)

atevossyan@yahoo.com

Резюме: В статье рассматривается возможность использования нейросетей Feed Forward для моделирования цен. В статье также представлены вопросы модельных перестановок в процессе машинного обучения. Рассматриваются три этапа генерации моделей в машинном обучении: обучение, валидация (калибровка) и тестирование (оценка), которые помогают избежать переустановок моделей.

Ключевые слова: искусственные нейронные сети, переустановка моделей, прогнозирование инфляции

Artificial neural networks in the process of price analysis

Tevosyan A.

International Scientific-Educational Center, NAS, RA (Armenia, Yerevan)

atevossyan@yahoo.com

Abstract: The article discusses the possibility of using Feed Forward Neural Networks for price modeling. The article also presents model overfitting issues in machine learning models. It discusses the three stages of model generation in machine learning: training (model fit), validation (calibration) and testing (evaluation), which help avoid overfitting.

Keywords: artificial neural networks, model overfitting, inflation forecasting

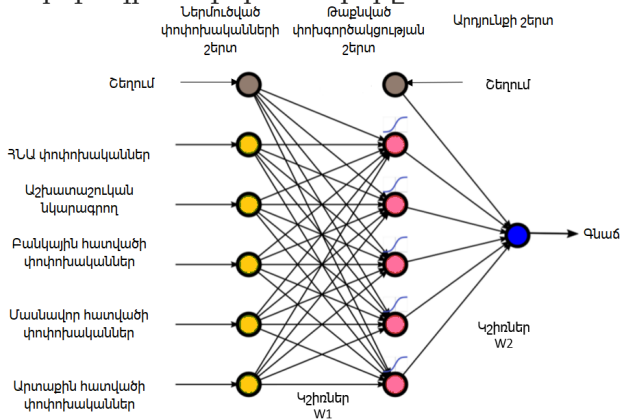
Մեքենաների ուսուցումն իրենից ներկայացնում է համակարգիչներին ինքնավար կերպով ուսուցանելու գիտություն՝ հիմնված իրական աշխարհի փոխազդեցությունների և մարդկային վարվելակերպի ուսումնասիրման, սերտման, կրկնօրինակման, և որ ամենակարևորն է՝ որակական բարելավումը և օպտիմիալացումը ապահովող ալգորիթմների վրա: Մեքենաների ուսուցումը կատարվում է բազմաթիվ մոդելներով: Մույն հոդվածում մենք կխոսենք մեքենաների ուսուցման թերևս ամենահզոր մոտեցման՝ նեյրոնային ցանցերի մասին: Նեյրոնային ցանցերը սկզբնազգծված են եղել մարդկային մտածողության կրկնապատկերով: Նեյրոնային ցանցերը հանդիսանում են ինքնաուսուցանվող համակարգեր: Ինքնաուսուցման ընթացքում, ունումնասիրելով և սերտելով բավականին մեծ ծավալի տեղեկատվություն, ցանցերը կարողանում են բացահայտել մուտքային և ելքային տվյալների միջև եղած թաքնված ամենաբարդ և հաճախ

նաև թաքնված կախվածությունները, կատարել ընդհանրացումներ և գնահատումներ: Հենց նեյրոնային ցանցերի մոդելների միջոցով է հնարավոր եղել հիմնել արհեստական բանականության գաղափարը:

Արհեստական նեյրոնային ցանցերի տեսակը, որը առավելապես օգտագործվում է գների մակարդի մոդելավորման համար, նախասնույց արհեստական նեյրոնային ցանցերն են (Feed Forward Neural Networks կամ FFANN), որոնք օգտագործելով բազմաշերտ ընկալիչներ, ի վիճակի են ընկալել անսահմանափակ քանակի շարունակական ֆունկցիաներ և մոտարկել դրանք՝ վերածելով ընդհանրական արդյունքների:

Նեյրոնային ցանցերը կարող են նկարագրվել որպես կիսապարամետրիկ մոդելներ: Մի կողմից, մոդելի կառուցվածքը և կապերը նախանշվում են ցանցային կառուցվածքի միջոցով, մյուս կողմից, սակայն, ցանցերը իրենք են որոշում, թե մուտքագրված տվյալ-

ների պայմաններում n թ կապերն են կարևոր և n թ կոնկրետ պարամետրերը:



Գծապատկեր 1: Նախասնույց արհեստական նեյրոնային ցանցերի (Feed Forward Neural Networks կամ FFANN) գրաֆիկական ներկայացումը:

Ցանցերի ունակությունը առավելապես բարելավվում է, երբ նախասնույց տվյալների քանակը աճում է:

Առաջնային փուլում ներմուծված տվյալները քայլ-առ-քայլ փոխգործակցության մեջ են մտնում, և յուրաքանչյուր հանգույց (Գծապատկեր 1, դեղին հանգույցները) իրենից ներկայացնում է համակցված փոխգործունեության ազդանշան մեկմեկու՝ ոչ գծային ֆունկցիաների միջոցով: Այնուհետև, հաջորդ փուլում, փոխգործակցությունը ավելի է խորանում և բարելավվում, երբ սկսում են ձևավորվել ներմուծված տվյալների կշիռները և հարաբերական պարամետրերը: Այս փուլում գործածվում է ոչ բացահայտ փոխգործակցության շերտը (Գծապատկեր 1, վարդագույն հանգույցները): Թաքնված հանգույցները իրենցից ներկայացնում են բարելավման կամ ակտիվացման ֆունկցիաներ (activation function), որոնք սերտելով ցանցային փոխկապակցվածության հատկանիշները և հարմարվելով բազմաթիվ վարքագծային դրսևորումների նվազագույն շեղումներով նմանակում են վերջնական արդյունքը, այս սեպքում՝ գնաճը (Գծապատկեր 1, կապույտ հանգույցը): Ընդհանրապես, ոչ գծային հարաբերությունները մոդելավորող ակտիվացման ֆունկցիաները ունեն լոգիստիկ կամ սիգմոիդ տեսք: Հաճախ նաև, երբ կարիք է լինում երկու կամ ավելի դասերի դասակարգային մոդելներ զարգացնելու (classes of classification models) օգտագործվում է Ֆրիդմանի սոֆթմաքս

(softmax) ֆունկցիան (Ֆրիդման և այլք, 2009):

Ինչպես տեսանելի է Գծապատկեր 1-ից, W_1 կշիռների մատրիցը ներմուծված տվյալները համակցում է ցանցի թաքնված փոխգործակցության շերտի հետ, որտեղից և ձևավորվում են հատկանիշները, որոնք W_2 կշիռների մատրիցի միջոցով ձևավորում է ներծին փոփոխականը՝ գնաճը: Հանգույցերի ակտիվացիոն ֆունկցիայի գործակիցները հենց կշիռների մատրիցի տարրերն են: Այսինքն, i -րդ հանգույցի k -րդ շերտի համար ակտիվացիան ֆունկցիայի մուտքային արժեք է i -րդ հանգույցի $(k-1)$ շերտի կշիռների մատրիցի գործակիցը: Գծապատկեր 1-ի պարզ բանաձևային տեղքը կլինի հետևյալը՝

$$LG(X \cdot W) \equiv \frac{1}{1+e^{-x \cdot w}} \in (0, 1),$$

$$Y = LG(X \cdot W_1^T) \cdot W_2^T \equiv h(X, \beta),$$

Որտեղ $X \cdot W$ իրենից ն ներկայացնում է X հատկանիշի և կշիռների մատրիցի արտադրյալը: Նշենք, որ լոգիստիկ ռեգրեսիան հնարավորություն է տալիս ներմուծել անփոփոգ գործակից (շեղում) ցանկացած շերտում:

Սովորաբար ազատ պարամետրերի քանակը որոշվում է թաքնված շերտերի քանակով և յուրաքանչյուր շերտում հանգույցների թվով: Ցանցերն, որոնք ունենում են ավելի քան մեկ թաքնված շերտ, կոչվում են խորը, և հենց այստեղից էլ խորը ուսուցում (deep learning) (ԼեՔուն և այլք (2015)¹: Գուդֆելլո և այլք (2016)²): Չնայած որ սովորաբար ակնկալվում, որ ազատ պարամետրերի աճի հետ մեկտեղ մոդելը ավելի լավ որոյունք կտա, մեծ է հավանականությունը, որ սա կրերի նաև մոդելի գերհարմարմարեցման (overfitting): Գերհարմարեցումը արհեստական բանականության և մեքենաների ուսուցման մեջ ամենախրական և վտանգավոր խնդիրներից է: Շատ դեպքերում ալգորիթմը չափից լավ է հարմարվում տվյալներին՝ առավելապես կենտրոնանալով մոդելի աղմուկի (noise), այլ ոչ թե ազդանշաններ վրա: Սա շատ հաճախ

¹ LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature, 521(7553):436–444.

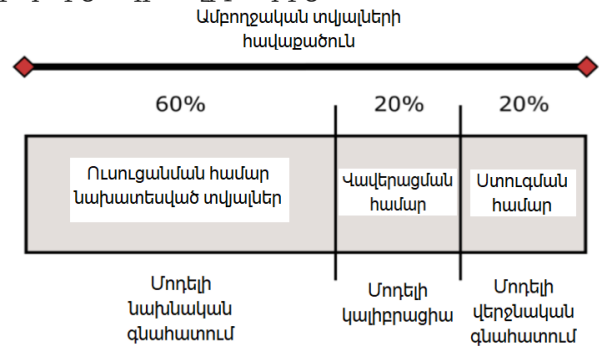
² Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep Learning. MIT Press.

պատահում է այն դեպքում երբ տվյալները սակավ են, իսկ մոդելը չափից առավել բարդ հարուստ հանգույցներով: Արդյունքում ունենում ենք ցանցային ալգորիթմ, որը թեև կատարելապես հարմարվել է նոր տվյալներին, բայց թույլ է և չի կարողանում ճշգրիտ կանխատեսումներ կատարել՝ հանգեցնելով ցածր որակի և իրականությանը չհամապատասխանող արդյունքի: Սովորաբար, այս դեպքերում, հարկ է մեքենայի ուսուցումը սկսել համեմատաբար պարզ և փոքր քայլերով, մոդելը փորձարկել վավերացման համար առանձնացրած տվյալների վրա և կատարել խաչաձև վավերացումներ (validation / cross-validation): Սովորաբար, մեդելների գերահամարեցման խնդրից խուսափելու համար, խորհուրդ է տրվում հետևել մոդելների փորձարկման և վավերացման երեք քայլերին: Մրանախապայան է ամբողջական թվային տվյալների բաջանումը երեք մասի՝ ուսուցման, վավերացման և ստուգման՝ խստագույնս պահպանելով պատահական ընտրանքի սկզբունքը բոլոր երեք տվյալին ենթախմբերի համար: Հարկ է նշել սակայն, որ բոլոր երեք խմբերում բավարար քանակի դիտարկումներ ունենալու համար ամբողջական թվային տվյալների հավաքածուն/բազան պետք է լինի բավարար մեծ:

Մոդելների նմանատիպ վերապատրաստման եռակողմ ցիկլը՝ մոդելի ձեւավորումը, հետագա վավերացումը եւ փորձարկումը միայն կարող է ապահովել մոդելի վստահելի ընդհարացնող հատկություններ:

Այսպիսով մոդելի օպտիմալ դիզայնը կախված է տվյալների շտեմարանի չափից,

հարցի բնույթից և շատ հաճախ մոդելավորողի փորձից և գիտելիքներից:



Գծապատկեր 2: Մեքենաների ուսուցման մոդելների երեք փուլերի սխեմատիկ ներկայացում՝ ուսուցում (մոդելի համապատասխանեցումը տվյալներին), վավերացում (ստուգում) եւ փորձարկում (զնահատում):

Օգտագործված գրականության ցանկ

1. LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature, 521(7553):436–444.
2. Glass E., Big data in central banks: 2016 survey, Central Banking Journal 27 (2) November 2016, last modified November 7, 2016, www.centralbanking.com/2474825 2.
3. Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep Learning. MIT Press.
4. Chiranjit Chakraborty and Andreas Joseph, Machine learning at central banks, September 2017
5. Thomas R. Cook and Aaron Smalter Hall, Macroeconomic Indicator Forecasting with Deep Neural Networks, September 2017.
6. Koop G., Onorante L., Macroeconomic Nowcasting Using Google Probabilities, August 2013.
7. Bholat, D. (2015). Big data and central banks. Bank of England Quarterly Bulletin, 55(1):86–93.
8. Einav, L. and Levin, J. D. (2013). The data revolution and economic analysis. Working Paper 19035, National Bureau of Economic Research.