**[[1]](#endnote-1)Моделиране на изкуствен интелект с помощта на невронни мрежи в геотехниката**

Проф. дтн Николай Николаев,

Проф д-р Кръстю Дерменджиев

**Резюме**

През последните десетилетия моделирането на изкуствен интелект /AI/ навлиза бурно и в областта на геотехниката и геотехническото строителство. Един от методите на това моделиране се осъществява чрез приложението на изкуствени невронни мрежи /ANN/.

На основата на литературни източници публикувани в диапазона на 1992 – 2024г. в този доклад са представени редица начини на използване на изкуствените невронни мрежи и проблемите при тяхното създаване за целите на геотехниката. Докладът има за задача да набележи само част от възможносттите на приложение на /AI/ в цитираната по-горе област.

Поради големият брой на използваните литературни източници те не са надлежно цитирани. Отбелязани са само имената на авторите, които са създали свои собствени модели.

Авторите нямат никакви претенции за собствени приноси в представеният доклад.

**Въведение**

Изкуствените невронни мрежи /ANN/ са форма на приложението им за моделиранe на изкуствения интелект / AI /, която се опитва да имитира функцията на човешкия мозък и нервната система. За реализиране на поставената цел е необходимо машината /компютър или специализирано създадено устройство/ да бъде предварително обучено чрез внедрявянето в него на определен набор от примерни опитни данни. Този процес се отбелязва като машинно обучение /ML/. Обучението се извършва в два начални стадия: плитко /начално/ и дълбоко- последващо. Първото обучение се характеризира с използването на относително малък брой данни, а при второто се проверява и коригира обучението със сравнително голям брой опитни данни.Плиткото обучение има за задача да запазнае машината с характера на задачата, като тя се запознае с финните-прецизни функционални връзки между данните, дори ако основните връзки не са съвсем известни и физическия генезис на процеса не е известен и труден за обяснение.

 Този характер на действие на ANN мрежите е в контраст с повечето емпирични и статистически методи, които се нуждаят от предварителни указания за характера и естеството на съществуващите връзки мужду данните.

 Това характерно свойство на моделирането с ANN е много подходящо за прогнозирането на поведението на скалната среда при разясняване на протичането на различни процеси в геотехниката с неизяснен характер

**Преглед на изкуствените невронни мрежи /ANN/**

Много автори са описали структурните особености на ANN. Един от тях е Eausett 1994, чийто труд от 429 стр., може да бъде изтеглен изцяло от интернет.

Описанията на ANN водят до различни видове набори от изкуствени неврони наречени „ Обработващи елементи „ /РЕ/, „възли“ или „единици“. Те са разпространени в отделни слоеве като образуват т.нар.“многослойни парцептрони“ /MLP/, които са най-често използваните ANN в геотехническото инженерство.

Най характерното и често пъти изпозвано разпределение е това при което са налични следните видове слоеве: входящ слой; изходящ слой и един или повече междинни наречени още „скрити слоеве“ в които са полегнали и така наречените „скрити възли“.



Фиг.1а. Структурна архитектура на ANN .

Всеки обработващ елемент в конкретния слой е напълно или частично свързан с много други обработващи елементи, чрез т. нар.тегловни връзки. Скаларните тегла определят връзката между взаимно-свързаните неврони. Нулевото тегло означава липсата на връзка между два неврона, а отрицателното тегло – забранена връзка. Тегловните стойности се сумират /със своите знаци/ и се заместват в така наречената предавателна функция /тава е или „ „ – логическата сиглоида или хиперболичната „ „, тангенс функция/, за получаване на изходните данни от обработката на елементите. За възела j този процес на сумиране в първото или второто уравнение е изобразен графично на фиг.1в.



Фиг.1в. Протичане на процесите

 -сумиране ;

 -трансфер ;

Където: -нивото на активиране на възела;

 – теглото на връзката между възли j и i ;

 – входът от възел I /I = 1,2,……..n /

 – отклонението за възела j ;

 - изходът на възела j ;

 - трансферната функция.

Разпространението на информацията в MLP започва от входния слой, където са представени входните данни. Тези данни се претеглят и се приемат от всеки възел в следващия слой. Претеглените данни след това се сумират и се предават чрез трансферната функция за получаване на възловия изход. Невронната мрежа коригира своите тегла съобразно представените обучаващи данни, като се използва правилото за обучениие до тогава, докато не се получат набори от тегла, които да определят изобразяването на входните и изходни данни с най-малката възможна грешка. Именно този процес е известен с термина обучение или тренировка.

 Допълнителното обучение е два вида: контролирано и неконтролирано. При първият вид са налице събрани от литературни източници входни и изходни данни за даден скално земен масив от определен-ограничен район на минен или строителен обект, така че е налице най-малкото един желан по стойност изход. Полученият изход от мрежата се сравнява с желания изход като се изчислява получената грешка. При наличие на по-голяма грешка се коригират приетите модели.

 Задачата има и обратно решение. Тръгва се от допустимата / минимална грешка/, минава се към прогнозиранепна вида на трансферната функция и към коригиране на невронната мрежа.

 При неконтролирано обучение на мрежата /машината/, се подават само наличните опитни или литературно оповестени данни за определени характеристики на минния масив, които са свързани с търсените неизвестни изходни данни, / липсват така наречените желани данни/. В този случай мрежата показва само входните данни. Идеята при това обучение е да се направи една класификация на входните данни отразяваща сегрегацията на отделни групи от данни със сходни структурни или физически признаци.

 Изкуствената невронна мрежа може да се категоризира въз основа на два основни признака: 1. Използваното правило за обучение (ML); 2. Връзката между обработващите елементи.

 От представеното по-горе става ясно, че в зависимост от връзките между обработващите елементи ANN, може да се раздели според начина на търсене в мрежите на търсене с права и търсене с обратна връзка.

 Философията за моделиране на ANN е подобна на редица конвенционални статистически модели. И при двата вида модели стремежът е да се намери връзката между събраните /налични/ входни данни на модела и съответния изход. Тоест търсене по аналогичен начин на връзката между вероятностния модел и най близкият му статистически аналог.

 Ако се предположи, че са налице набор от „х“ стойности от дадено събитие, на което съответстват „у“ стойности на решенията в двумерното пространство, целта е да се намери неизвестната функция , която свързва входните резултати „ Х“ с изходните „У“.

 При линеен регресионен модел функциятаоже да се получи по сравнително прост начин чрез нагаждане на наклона на правата и отрязъка на правата „β“, така че грешката между фактическите изходни данни и тези от правата да се минимизират. В този случай ANN може да се оформи като прост линеен регресионен модел с един вход и линейна предавателна функция.

 ANN може да коригира многократно своите тегла, като обработва и изменя входните и изходните даннни минимизирайки функцията на грешките.

 Ако връзката между „Х“ и „У“ е нелинейна, регресионнният анализ може да се приложи успешно само ако съществуват предварителни познания за характера на нелинейността /характерни особености на подходящата крива/.

 В моделът ANN видът за нелинейност може да се променя, чрез промяната на трансферната функция и броят на скритите /междинните/ възли на слоя. При сложни и силно нелинейни случаи, традиционният регресионен анализ е неадекватен. В такива случаи може да се прибегне до смяна на: трансферната функция; изменение на мрежовата структура; промяна на броят на скритите слоеве и броят на възлите във всеки слой.

**Проблеми при моделирането на изкуствените невронни мрежи / ANN /**

 Подобряването на производителността на ANN зависи отдобрата и систематична обработка на входните данни. От особена важност е да се обърне внимание на основните фактори определящи адекватността на входните данни, разпределението им по определени характерни особености, изборът на подходяща архитектура на модела, внимателен подбор на някои вътрешни параметри, които контролират методът на оптимизация. Всички отбелязани фактори ще бъдат обект на анализ по-нататък.

**Определяне на входовете на модела**

Основна стъпка при разработването на ANN моделите е изборът на входните променливи, които имат най значимо въздействие върху производителността му. Добавянето на възможно най-голям брой входни данни обикновено увеличава мрежата, което води до намаляване на скордстта на обработка и намалява нейната ефективност.

 В специалната литература са предложени редица техники за подпомагане при избора на входните данни. Подходът, който най-често се използва в областта на геотехниката е: изборът на входни данни да се съсредоточи само върху тези, които съгласно априорните знания са най-адекватни за случая. Друг подход е да се обучат много невронни мрежови модели с различни комбинации от набори на входни данни и да се избере този модел, който осигурява най добри показатели /резултати/.

 Използва се също и многостъпкова тихника описана от Maier и Dandu (2000год. ), при която се обучават отделни мрежи, като всяка от тях използва само една от наличните променливи за вход на модела. След това мрежата с най-добри показатели се запазва. По-нататък към нея се прибавят следващите по степен на достоверност модели, до получаването на допустимото отклонение за дадената задача.

Друг полезен метод е използването на генетичен алгоритъм за търсене на най-добрият набор от входни данни. Избраната съобразно този алгоритъм мрежа се обучава и използва за класиране на различните подмножества от входни данни. Наборът от входни променливи получава своята степен на годност от оценената грешка на модела.

 Алгаритъмът за адаптивно сплайн-моделиране на оптични данни /представен с абревиатурата A.S.MOD /, също се сочи като една от полезните техники, която може да се използва за избора на „пестеливи“ невронни мрежи, чрез автоматично избиране на комбинация от входни променливи, която ще има най-значително въздействие върху резултатите.

 Потенциален недостатък на споменатите подходие, че те са базирани на предварително избран модел. За да се елиминира този недостатък може да се използват зависимости от рода на линейни и нелинейни корелации, такива, които да осигуряват взаимната информация за подбора на подходящите входни данни

В други литературни източници определянето на това дали даден модел е адекватен не зависи само от избора на входните данни, но и от подбраната структура и калибровка на моделите. За да се елиминира този недостатък може да се използват зависимости от рода на нелинейни корелации, които осигуряват взаимната информация за подбора на подходящите входни данни .

**Опитни данни**

 Контролираните опитни данни за ANN са подобни на конвенционалните статистически модели. Параметрите на модела, / например теглата на връзките / се коригират при калибриране на модела – при обучението / ML /. При това се минимизират грешката между изходните данни и съответните измерени стойности за конкретният набор от данни – наборът за обучение. ANN се праедставят най добре, когато не екстраполират извън обхвата на данните за калибриране, за разлика от конвенционалните статистически модели.

 Наличните данни се разделят на две подмножества. Данните от първото подмножество, обикновено около две трети от данните участва в обучението, а второто, останалите данни за оценка на адекватността на модела в окончанелния му вид. Модификация на горния подход е формиране на три групи данни. За обучение, тестване и крайна проверка на модела. Този подход предпазва модела от „преобучение“ чрез използване на междинното тестване.

 Такъв начин на провеждане на разпространението на опитните данни се използва при прогнозирането на поведението на монтираните плитки фундаменти. Изследванията са показали, це най-доброто разпределение е 20% от данните да участват в обучението, 60% в тестването и 20% в проверката. В случаите, когато броят на данните не е голям се използва т.нар. метод на изключването, състоящ се в задържане на малка част от данните за валидиране и цялата останала част за обучение.

 Проучванията на друга група изследователи обаче установяват, че начинът по който се разделят данните, може да окаже значително влияние върху получените резултати. Тъй като определения ANN модел понякога среща затруднения,когато се налага екстраполация извън обхвата на калибрираните данни. За да се разработи един подобрен модел на ANN, трябва да бъдат включени и екстремните точки в наличните данни, които са били изключени при калибрирането. Този подход осигурява способността на модела да работи не само в интерполационния но и векстраполационния интервал.

 Може да се отбележи и факта, че статистическите свойства на различните под-множества от данни, като например средното и средно стандартното отклонение на данните, трябва да бъдат сходни с тези за валидиране, обучение и тестване, за да се гарантира, че всяко подмножество представлява една и съща статистическа популация. Ако това не е така, може трудно да се прецени валидността на ANN модела.

 Базирайки се на това предположение, в няколко изследвания са били използвани специални методи за да се гарантира, че използваните данни притежават и посочените статистически свойства, което е потвърдило и стратегията за разделянето на данните в подмножества.

 Изискванията към входно-изходните данни, начинът на тяхното разделяне и групиране и включването им в процесите на обучение,тестване и крайната проверка оказва голямо влияние върху ANN. Във връзка с това през 2002 г. в литературата се появяват данни за подходящ подход за разделяне на данните в подмножества отговарящи на всички изисквания [Boudek & ad,2002]. За целта авторите са използвали генетичен алгоритъм свеждащ до минимум разликите между средните стойности и стандартните отклонения. Известен пропуск в посочения метод е факта, че не се дават насоки за определяне на отделните подмножества, наречени „ клъстери“.

**Предварителна обработка на данните**

 След разделяне на данните в подмножества се извършва подходяща обработка. Тя е необходима за гарантиране равностойното внимание към всички участващи променливи в процеса на обучение.Тази обработка може да се състои под формата на мащабиране, нормализиране и трансформация на данните.Мащабирането на изходните данни е от съществено значение, тъй като данните данните трябва да бъдат с границите на трансферните функции, използвани в изходния слой / например между „-1“ и „+1“, ако за това е приета тангенсовата трансферна функция или от 0,0 до +1, при сигмоидалната трансферна функция/.

 Трансформацията на входните данни става в някои известни форми /например линейни, логаритмични, експоненциални и др./, което при подходящ избор на трансформация може да доведе до подобряване производителността на модела на ANN.

**Строеж на моделите**

 Определянето на строежа на мрежите е една от най трудните задачи при разработването на ANN моделите.Това изисква определянето на оптималния брой на слоевете и броят на връзките в тях.

 До момента все още няма унифициран подход за създаванено на оптимална архитектура за ANN моделите. При съществуващите съвременни възможности е необходимо това да се постигне на базата на известен опит у изследователя и съобразяване с характерните особености на задачата.

Началото изисква определяне на броя на слоевете и лежащите в тях брой на възлите за всеки от тях. Освен входните слоеве се изисква и уточняване на броя и състава на скритите /междинни/ слоеве. Характерната особеност тук се състои в това, че именно броя на скритите слоеве определят гъвкавостта /нагаждането/ на моделирането, т.е. пригаждането на ANN към определени условия. Добре подбраният брой на тези слоеве дава възможност за оптимално моделиране и формиране на по-сложни функции.

Изследователите предлагат различни решения по този въпрос. Например според Faeber и колектив наличието на два скрити слоя е напълно достатъчно за решаване на повечето задачи в областта на геотехниката. Това твърдение е подкрепено в публикувания от тях труд с достатъчно практическо доказателство.

Според Chester първият скрит слой се използва за извличане на глобалните характеристики на съставения модел за обучение. Мастерс заявява, че използването на повече от един скрит слой често пъти забавя драстично процеса на обучение и увеличава шанса да се попадне в „ капана“ на локалните минимуми.

Броят на възлите във входния и изходящия слоеве е ограничено съответно от броя на входовете и изходите на модела. Няма директен и точен начин за определяне на най-подходящия /оптимален/ брой на възлите във всеки скрит слой. Най-често се използва процедурата „проба – грешка“, която в преобладаващите случаи се прилага при моделирането на задачи в областта на геотехниката. Същият подход се използва и при определяне броя на връзките на възлите в скрития слой.

Окончателното заключение съставено от мнението на голям брой изследователи е „Запазването на скритите слоеве до възможния минимум, който осигурява задоволителна производителност на модела“, което води до: намаляване на изчислителното време необходимо при обучението / ML /; помага на мрежата ва осигури по представителна генерализация на проблема; помага да се избегне прекомерното насищане на модела, което често пъти става повод за объркване на връзките; позволява обучената мрежа да се анализира по-лесно.

За мрежи с наличието на само единскрит слой има редица практични правила за получаване на най-добрия брой връзки в скрития слой. Един от подходите е : „броят на скритите възли да бъде между средната стойност на връзките във входния и изходния слоеве.

Друг подход фиксира горната граница на броя на скритите възли в еднослойната мрежа по правилото /2⌡+1/, където: ⌡е броят на връзките във входа.

Според мнението на достатъчен брой изследователи /обсъдено на специализирани конференции/, най-добрият подход е този предложен от Nawary и колектив е да се започне с малък брой скрити възли и да се премине към слабо, последователно увеличение на възлите, докато се постигне значително подобрение на модела. Част от китайските и специалистите от Малайзия и Сингапур предлагат друг начин за определяне на оптималния брой на скритите връзки.Броят им да се приравни на броя на тренировъчните проби.

Maier и колектив /2000г./, твърди, че начина да се попречи на мрежата „да научи“, уникални, „строго персонални“ характеристики, което намалява обобщаващият и вид, е да се използват оптимален брой данни. Приемането на голям брой примери, от една страна се смята, че е най добрия подход, но от друга това води до разводняване на подадената информация и до голямо разсейване на данните.

Съществуват и няколко практически правила, предложени в специалната литература отнасящи се до свързване на броя на тренировъчните проби с броя на теглата /т.е. връзките им/, а именно: броят на теглата /връзките/ да не надвишава броя на обучителните проби.

Канадски изследователи предлагат, че правилното съотношение между обучителните проби и броят на теглата трябва да бъде „2“.

Съществуват и редица систематични подходи за автоматично получаване на стандартна мрежова архитектура. Такъв например е „Адаптивният метод“ предложен от Glabosi & Sidarta, който се базира на приемането за първоначално по-малък брой връзки в скритите слоеве. По време на обучението и когато мрежата достигне капацитета си се генерират нови връзки в скритите слоеве. Обучението продължава и с новите възли за да се оформят и новите им тегла. Тези стъпки се повтарят и в случай на нужда /т.е. ако полученият резултат не осигурява добро съвпадение с крайните входно-изходни данни/ се прибавят нови скрити връзки.

Друга използвана техника е „методът на изрязване“ предложен от Карник. Той се характеризира с приемането на първоначална мрежа с голям брой на скритите възли, като в процеса на обучение излишните възли се премахват.

Генетичните алгоритми представляват друга еволюционна алтернатива за определяне на подходящата архитектура на ANN. Тя се използва успешно при много ситуации, но описването и е обемно и не е обект на този доклад. Известен е и методът „Адаптивно сплайн моделиране“, A.S.MOD, който представлява автоматичен метод за получаване на оптимална архитектура с помощта на т.нар.В-сплайн мрежа.

Fahman & Lebier предлагат, така наречената Casade- корелация, което също е автоматичен метод реализиращ оптимална мрежова архитектура. Характеризира се с това, че в първоначалната мрежа има директна връзка между входния и изходния слой, без скрит слой. Скритите слоеве и възли се добавят последователно на случаен принцип / един или няколко едновременно/. Новите скрити връзки получават връзки, както с оригиналните входове и изходи, така и с предишно монтираната част от скритите възли. Процесът спира, когато модела покаже необходимото подобрение. Конструктивният характер на този метод се изразява в предимството от формиране на по-малки оптимални мрежи, подобрена способност за обобщение на задачата и на по-висока производителност.

**Оптимизация на мрежовия модел**

Процесът на оптимизация на теглата на връзките е известен с термина „ обучение „. Този процес е еквивалентен на оценката на параметрите в конвенционалните статистически модели. Целта е да се намери глобално решение на една, в повечето случаи, нелинеина оптимизация.

Глобалната оптимизация може да се реализира чрез използването на методите за извеждането или приемането на известен генетичен алгоритъм, при оптималната комбинация от „тегло на подаване напред“ /MLP/ в невронните мрежи, което е алгоритъм основаващ се на градиент от първи ред.

Предимствата на тези методи е, че имат способността да избягват локалните минимуми на реализираните грешки и да осигурят оптимални или близки до оптималните решения на задачата. Като цяло може да се обобщи твърдението, че видът на проблема диктува ефективния модел и съответния алгоритъм и осигурява добра скорост на обучение.

**Критерии за стопиране на тренировъчния процес**

За определяне на момента на стопиране на тренировъчния процес може да се използват много различни подходи. Обикновено след изготвянето на подходящ брой записи на резултати от тренировъчния процес, при който крайните резултати осигуряват грешка с достатъчно малка стойност, когато настъпват много малки промени във величината на грешката или липсват каквито и да са изменения, /т.е. няма промяна/ в получената грешка, процесът на обучение се прекратява. По време на обучението се осигуряват записи от получените резултати при измененията на броя на слоевете и възлите.

Съществуват и различни критерии за определяне на момента на стопиране на процеса на обучение. Така например може да се използват „ Информационните критерии на Баеснан и Акаике, но във всички случаи основния критерий за стопиране си остава „величината на крайната грешка, реализирана при прогнозиране на крайния резултат.

**Валидиране /проверка/ на модела**

След приключване на обучението /ML/, получените резултати се подлагат на валидиране. Целта на тази проверка е да се установи дали полученото решение отговаря на обсега на заложените данни за обучение и порядъка на точност на взаимовръзката на входните и изходните данни.

Подходите за реализирането на тази важна проверка, посочени в различни литературни източници е определяне ефективността на получената „производителност на обучителния процес“ и се реализира с помощта на остатъка от опитните данни неучаствали в процеса на обучение. Ако се окаже, че и при този набор от данни, грешката на модела е в допустимите граници, моделът се счита за адекватен и е с възможности да осигурява надежни резултати от опитните данни.

Коефициентът на корелация „r“, средноквадратичната грешка /RMSE/ и абсолютната грешка /MAE/ са основните критерии за определяне „производителността“ на получения прогнозен модел за ANNs.

За обсега на степента на производителност изследователя в областта на геотехниката Shmith предлага следните степени за норми на производителност: при I r I ≥ 0,8 сигурна корелация между двата набора от променливи; при 0,2 ≤ I r I ≤ 0,8 наличие на известна корелация; : при I r I ≤ 0,2 налична е много слаба корелация.

В корелационния анализ, в областта на техниката, се предлагат други степени за определяне на годността на дадена регресия определена на базата на опитни данни, а именно: при I r I ≥ 0,7 съществува добра корелация; при I r I ≤ 0,7 корелацията не е подходяща за оценка на регресия използваща се в техническото строителство. При отговорни инженерни съоръжения се изисква понякога I r I ≥ 0,9, а дори и I r I ≥ 0,95.

**Предимства на ANN моделите**

В практиктиката на геотехническото инженерство, в някои случаи възникват сложни проблеми свързани с физическата им същност. В такива случаи ANN моделирането осигурява предимства пред конвенционалните емпирични или статистически изследвания. Така например при повечето математически модели липсата на известна физическа постановка на задачата бива преодолявана чрез известно опростяване на проблема, или чрез включване на определени субективни допускания. В други случаи решенията се базират на предварително приета структура на модела, което води до оптималното решение.

За разлика от посочените по-горе опростени решения ANN моделите използват подход при който наличните входни данни се обучават по отделни двойки входно-изходни данни, при което автоматично се отчитат структурните и други особености на континиума, в случая земно-скалния масив. Освен това резултатите от добре обучените и валидирани модели на ANN може да бъдат актуализирани чрез представяне на нови примери за обучение, когато са налице нови опитни данни. По-дулу на фиг.2 е представен примерен модел на учени от Япония използван за определяне слягането на повърхността под тунели в зависимост от дълбочината.



Фиг.2.

**Литература:**

1. [↑](#endnote-ref-1)