**ԵՐԵՎԱՆԻ ՊԵՏԱԿԱՆ ՀԱՄԱԼՍԱՐԱՆ**

**ՏՆՏԵՍԱԳԻՏՈՒԹՅԱՆ ԵՎ ԿԱՌԱՎԱՐՄԱՆ ՖԱԿՈՒԼՏԵՏ**

**Տնտեսագիտության մեջ մաթեմատիկական մոդելավորման ամբիոն**

**ՏՎՅԱԼՆԵՐԻ ԳԻՏՈՒԹՅՈՒՆԸ ԲԻԶՆԵՍՈՒՄ**

**ԿՐԹԱԿԱՆ ԾՐԱԳԻՐ**

**ԿՈՇԿԱՐՅԱՆ ՀՌԻՓՍԻՄԵ ՎԱՐԴԻՎԱՐԻ**

**ՄԱԳԻՍՏՐՈՍԱԿԱՆ ԹԵԶ**

**ԱՐՀԵՍՏԱԿԱՆ ԲԱՆԱԿԱՆՈՒԹՅԱՆ ԿԻՐԱՌՈՒԹՅՈՒՆԸ ԱՌՈՂՋԱՊԱՀՈՒԹՅԱՆ ՈԼՈՐՏՈՒՄ**

***«Տնտեսագիտություն» մասնագիտությամբ***

***Տնտեսագիտության մագիստրոսի որակավորման աստիճանի հայցման համար***

**ԵՐԵՎԱՆ 2020**

***Ուսանող` ———————————————————————————***

***Հռիփսիմե Կոշկարյան***

***Գիտական ղեկավար` ———————————————————————***

***տ. գ. դ․ պրոֆ. Ռ. Գևորգյան***

***«Թույլատրել պաշտպանության»***

***Ամբիոնի վարիչ` —————————————————————————***

***տեխ. գ. դ. պրոֆ. Ա. Առաքելյան***

*«\_\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_20\_\_\_թ.*

**ՀԱՄԱՌՈՏԱԳԻՐ**

***ԱՐՀԵՍՏԱԿԱՆ ԲԱՆԱԿԱՆՈՒԹՅԱՆ ԿԻՐԱՌՈՒԹՅՈՒՆԸ ԱՌՈՂՋԱՊԱՀՈՒԹՅԱՆ ՈԼՈՐՏՈՒՄ***

***MEDICAL IMAGE PROCESSING***

***ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В МЕДИЦИНЕ***

Վերջին տարիներին տեղեկատվական տեխնոլոգիաների աննախադեպ առաջընթացի, մեծ տվյալների բազաների հասանելի աղբյուրների առկայության և նեյրոնային ցանցերի ալգորիթմերի նկատմամբ մեծ հետաքրքրության զուգակցմամբ՝ խորը ուսուցումը (Deep Learning) մեծ տեմպերով զարգացում ապրեց: Արհեստական բանականությամբ լուծվում են մի շարք բնագավառների կարևորագույն խնդիրներ, այդպիսի բնագավառներից է բժշկությունը։

Այս աշխատանքում շեշտը հիմնականում դրվել է ոչ թե արհեստական բանականության նոր ալգորիթմների զարգացման վրա, այլ դրանց կիրառությամբ առողջապահության ոլորտում կարևորագույն խնդրի լուծմանը։ Մեքենայական ուսուցումը և արհեստական բանականությունը այնպիսի բնագավառների խնդիրներ կարող են կարող են լուծել, որտեղ գեներացվում են մեծ քանակությամբ տվյալներ և այդ տվյալներով հնարավոր է արժեքավոր մոդելներ կառուցել։ Բժշկությունը հենց այդ բնագավառներից մեկն է համարվում և արհեստականա բանականության գործիքները կարող են ավելի քան օգտակար լինել հենց այստեղ։

Աշխատանքը բաղկացած է երեք գլխից։ Հետազոտության նպատական է արհեստական բանականության կիրառմամբ կառուցել այնպիսի մոդել, որը կախտորոշի մարդու աչքի կուրության մակարդակը՝ շաքարային դիաբետով պայմանավորված։ Առաջին գլխում տրվելու է խնդրի նկարագիրը, նկարագրվելու են այն ուսումնասիրությունները, որոնք արվել են բժշկության մեջ արհեստական բանականության կիրառմամբ։ Երկրորդ գլխում նկարագրվելու է աշխատանքի բուն խնդիրը և տրվելու են այն մոդելների նկարագրությունները, որոնցով լուծվել է խնդիրը։ Երրորդ գլուխը ամփոփվելու է երկրորդ գլխի մոդելների արդյունքներով, այդ արդյունքերի համեմատություններով։

**ԲՈՎԱՆԴԱԿՈՒԹՅՈՒՆ**

# **ՆԵՐԱԾՈՒԹՅՈՒՆ**

Ներկայում աշխարհում աննախադեպ զարգացում է ապրում տվյալների գիտությունը։ Տվյալների գիտության տարբեր ճյուղերը գրեթե բոլոր բնագավարներում գտել են իրենց կիրառությունը։ Առողջապահության ոլորտը ևս աննմասն չի մնացել տվյալների գիտության կիրառությունից։ Տվյալների գիտությունը բժշկության մեջ արագացնում, հեշտացնում և զարգացնում է մի շարք հիվանդությունների ախտորոշումը, տվյալների գիտության միջոցով լուծվում են մարդու առողջությանը վերաբերող մի շարք կարևորագույն խնդիրներ, և սույն աշխատանքի էականությունը կայանում է հենց դրանում։

Բժշկության մեջ ախտորոշման խնդիրը նաև բիզնես խնդիր է համարվում։ Արհեստական բանականության կիրառությունը բիզնեսում հեշտացնում է բժիշկների և հիվանդների կյանքը՝ կրճատելով հիվանդության ախտորոշման ժամանակահատվածը և ծախսերը։ Առողջապահության ոլորտում արհեստական բանականության կիրառությունը տնտեսական աճի մեծ խթան կարող է լինել։ Ըստ «PwC» կանխատեսման առողջապահության ոլորտում արհեստական բանականության առաջխաղացումը կնպաստի համաշխարհային տնտեսության 15,7 տրլն դոլար արժողությամբ տնտեսական խթանմանը:

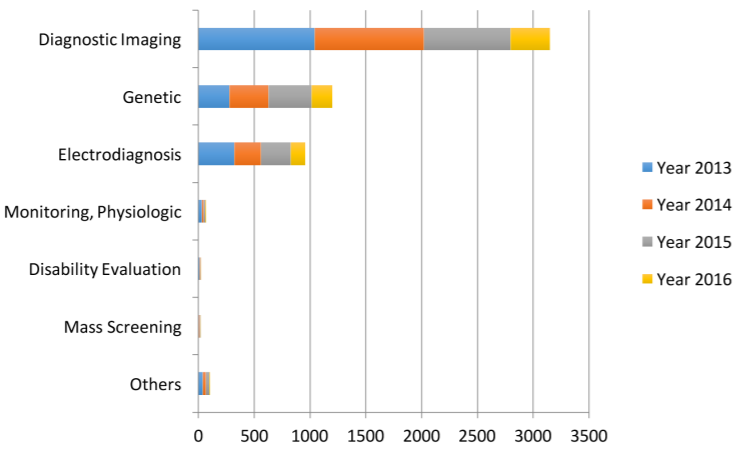
Արհեստական բանականությունը ներկայիս արագ փոփոխվող տնտեսության մեջ վերափոխումների, խթանման և մրցակցային առավելության հիմնական աղբյուրն է:

# **ԳԼՈՒԽ 1. ԱՐՀԵՍՏԱԿԱՆ ԲԱՆԱԿԱՆՈՒԹՅԱՆ, ԽՈՐԸ ՈՒՍՈՒՑՄԱՆ ԿԻՐԱՌՈՒԹՅՈՒՆՆԵՐԸ ԱՌՈՂՋԱՊԱՀՈՒԹՅԱՆ ՈԼՈՐՏՈՒՄ**

# **1.1 Արհեստական բանաակնության դերը առողջապահության ոլորտում**

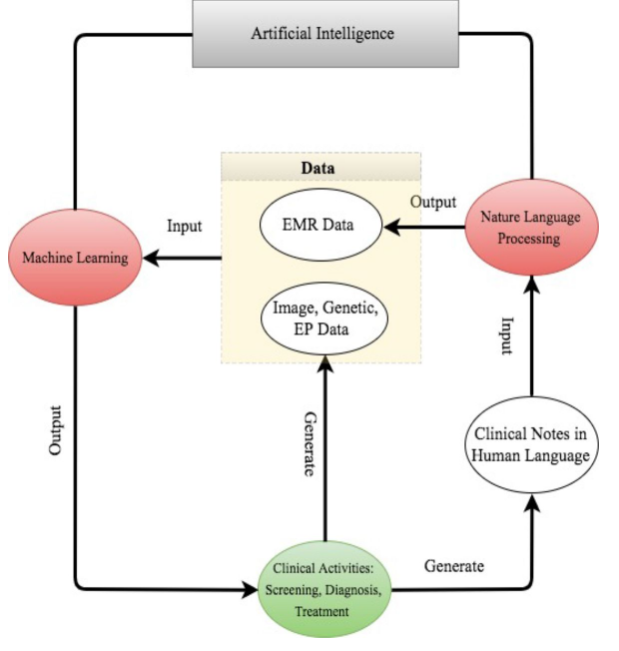
Արհեստական բանականությունը վերջերս հասել է մեծ հաջողությունների առողջապահության ոլորտում։ Արհեստական բանականությունը կարող է գեներացնել մի շարք բարդ ալգորիթմներ «մեծ տվյալներից» (Big data) անհրաժեշտ արդյունքներ ստանալու համար, և այնուհետև հնարավոր է այդ արդյունքները ներդնել պրակտիկ բժշկական մի շարք խնդիրների լուծման մեջ։ Արհեստական բանականությունը կարող է կիրառվել տարբեր բժշկական տվյալների համար, այդ թվում և կառուցվածքային, և ոչ կառուցվածքային (structured and unstructured data): Մեքենայական ուսուցման հիմանական ալգորիթմները, որ օգտագործվում են կառուցվածքային տվյալների համար ներառում են Support vector muchine (SVM)-ը և նեյրոնային ցանցերով խորը ուսուցումը, իսկ ոչ կառուցվածքային տվյալների համար հիմնական կիրառվող ալգորիթմը համարվում է Natural language processing-ը (NLP): Բարդ հիվանդությունների ախտորոշման ոլորտները, որոնց համար կիրառվում է արհեստական բանականությունը ներառում են քաղցկեղի ախտորոշումը, նյարդաբանությունը և սրտաբանությունը։

Տվյալները, որոնք օգտագործվում են ախտորոշման փուլում հիմնականում ներառում են բժշկական նկարները, գենետիկ ուսումնասիրության արդյունքները, էլէկտրոդիագնոզը, հաշմանդամության գնահատուման արդյունքները և այլն (Գծապատկեր 1.):



Գծապատկեր 1. Արհեստական բանականության մշակման բժշկական տվյալները ըստ տարիների

Ընդհանուր առմամբ արհեստական բանականության տեխնիկաները բժշկության մեջ բաժանվում են երկու խմբի՝ առաջին խումբը իր մեջ ներառում է կառուցվածքային տվյալների մշակում, ինչպիսին են օրինակ նկարները, իսկ երկրորդը Natural language processing-ի մեթոդներն են, որոնք գործ են ունենում ոչ կառուցվածքային տվյալների հետ, մասնավորապես ինֆորմացիա է ստացվում բժշկական տարբեր հոդվածներից հետազոտություններ իրականացնելու համար։



Գծապատկեր 2. Արհեստական բանականությամբ բժշկական տվյալների մշակման հոսքագիծ

**Դասական մեքենայական ուսուցման ալգորիթմները բժշկական տվյալների համար**

Արհեստական բանականության ալգորիթմների միջոցով հնարավոր է դառնում տվյալներից մեզ պետքական առանձնահատկություններ ստանալ։ Որպես բժշկական տվյալներ, կարող են օգտագործվել հիվանդների հատկությունները (patient's traits), և հիվանդների ախտորոշման այն հատկությունները, որոնք իրենցից հետաքրքրություն են ներկայացնում։ Հիվանդների հատկությունները ներառում են հիվանդի տաիրքը, սեռը, հիվանդության պատմությունը, ինչպես նաև ախտորոշման համար հատուկ տվյալներ, ինչպիսին են օրինակ ախտորոշման նկարները։ Մոդելը կառուցելու համար անհրաժեշտ են բացատրող և բացատրվող փոփոխականներ, եթե բժշկական տվյալներում բացատրվող՝ կանխատեսվող փոփոխականներ են համարվում հիվանդության ինդիկատորները, հիվանդի ապրելը տևողությունը, հիվանդության մակարդակը և այլն:

Դասական մեքենայական ուսուցման ճյուղավորումները՝ վերահսկվող (supervised learning) և ոչ վերահսկվող (unsupervised learning) ուսուցումները, կախված տվյալներից կիրառվում են հետազոտություններում։ Ոչ վեչահսկվող ուսուցման ժամանակ տվյալները չհատկանշված են (unlabeled), և այն հիմնականում օգտագործվում է հատկանիշների կանխատեսման համար, մինչդեռ վերահսկվող ուսուցման ժամանակ տվյալները հատկանշված են (labeled), և վերահսկվող ուսուցումը փորձում է կառուցել այնպիսի մոդելներ, որտեղ կապեր կհաստատվեն հիվանդների հատկանիշների և կանխատեսվող արդյունքների միջև։

Քլասթերացումը (clustering) և հիմնական բաղադրիչի վերլուծությունը (principal component analysis) ոչ վերահսկվող ուսուցման երկու հիմնական մեթոդներն են։ Քլասթերացման ժամանակ միանման հատկանիշները համախմբվում են միևնույն խմբում, առանց նախապես հատկանշված տվյալների։ Քլասթերինգի ալգորիթմները հիվանդների բաժանում է քլասթերների՝ մաքսիմալացնելով հիվանդների նմանությունը քլասթերների միջև և մինիմալացելով հեռավորությունը քլասթերներից դուրս։ Հայտնի քլասթերինգի մեթոդներն են «k-means» քլասթերինգը, «hierarchical» քլասթերինգը և «Gaussian mixture clustering» քլասթերինգը։ Հիմնական բաղադրիչի վերլուծությունը չափողականության կրճատման համար է, երբ հիվանդի հատկանիշը ներմուծվում է մեծ քանակի չափողականությամբ, ինչպես օրինակ մարդու գեների քանակը գեների ուսումնասիրության մեջ։ Հիմնական բաղադրիչի վերլուծության միջոցով մեծ քանակությամբ տվյալներից հնարավոր է լինում մի քանի հիմանական տվյալներ առանձնացնել, առանց տվյալների բազայի ինֆորմացիայի էական կորստի։ Երբեմն կատավում է հիմնական բաղադրիչի վերլուծություն՝ չափողականության կրճատման համար, այնուհետև նոր արվում է քլասթերինգ խմբավորումը։

Վերահսկվող ուսուցման ժամանակ փորձ է արվում համախմբված տարբեր հատկանիշների միջոցով ստանալ այնպիսի արդյունք, որը հնարավորինս մոտ կլինի ակնկալվող արդյունքին։ Այս դեպքում հստակ սահմանված են լինում օբյեկտի հատկանիշները, նպատակը այնպիսի մոդելի կառուցումն է, որը կունենա արդեն իսկ սահմանված հատկանիշներին մոտ արդյունք։ Օրինակի համար արդեն իսկ ախտորոշված հիվանդության տվյալներին կառուցել հիվանդությունը կանխորոշոխ մոդել։

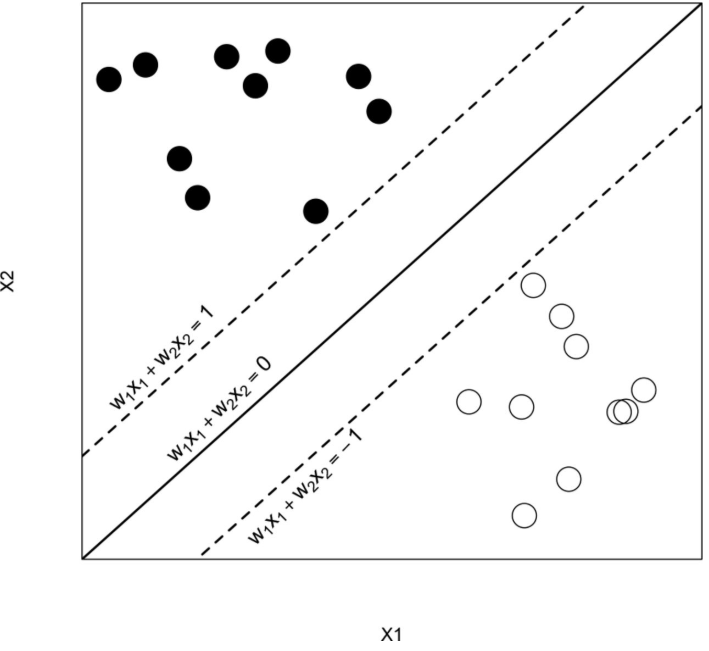
Եթդ համեմատենք վերահսկվող և ոչ վերահսկվող ալգորիթմները, ապա վերահսկվող ալգորիթմենրը տալիս են ավելի համապատասխան արդյունքներ, այդ իսկ պատճառով արհեստական բանականության կիրառությունները բժշկությունում։ Հիմնականում վերլուծությունների ժամանակ կիրառվում է ոչ վերահսկվող ուսւոցումը տվյալների չափողականության կրճատման համար, այնուհետև նոր կիրառվում են վերահսկվող ուսուցման ալգորիթմները, ինչը ավելի էֆեկտիվ է դարձնում ցանկալի արդյունքի հասնելը։ Վերահսկվող ուսուցման հիմնական մեթոդներն են ռեգրեսիաները, որոշման ծառերը (Decision trees), «KNN»-ը, «Պատահական անտառը»-ը (Random Forest), Support vector machine (SVM)-ը և նեյրոնային ցանցերը։

# **1.2 Մեքենայական ուսուցման մի քանի մոդելների կիրառությունը**

Support vector machine (SVM)-ը օգտագործվում է օբյեկտները խմբերի բաժանելու համար։ Եթե արդյունքը, որը ակնկալում են ստանալ SVM-ը կիրառելուց հետո նշանակենք , ապա օրինակի համար ցույց է տալիս հիվանդի առաջին կամ երկրորդ խմբին պատկանելը։ Եթե նշանակենք հիվանդի հատկությունը , ապա հիվանդի այս կամ այն խմբին պատկանելը կարող է բաժանվել որոշման գծի շնորհիվ (a decision boundary): Մաթեմատիկորեն կստացվի հետևյալ հավասարումը․

, որտեղ

-։ Եթե , հիվանդը դասակարգվում է առաջին խումբ, այս դեպքում մենք ենք դասակարգում, իսկ երբ, հիվանդը դասակարգվում է երկրորդ խումբ, կամ որ նույնն է ենք դասակարգում: Երբ հիվանդի դասակարգումը անորոշ է։

**

Գծապատկեր 3. SVM-ի արտապատկերում երկու փոփխականի համար

Հիմնական նպատակն է գտնել այնպիսի , որոնք հնարավորինս ճիշտ կդասակարգեն հիվանդին՝ նվազեցնելով դասակագման սխալանքը։ Լավագույն դասակարգման դեպքում և արժեքները պետք է համընկնեն, և բավական հեռու լինի 0-ից, որ դասերի միջև անորոշությունը հնարավորինս քիչ լինի։ Սա կարող է լինել այնպիսի ընտրությամբ, որ մոդելի «կորուստը» (loss) ձգտի մինիմումի։ Այնուհետև նոր հիվանդներին կարող ենք արդեն դասակարգել արդեն իսկ «վարժեցված» (trained) միջոցով։

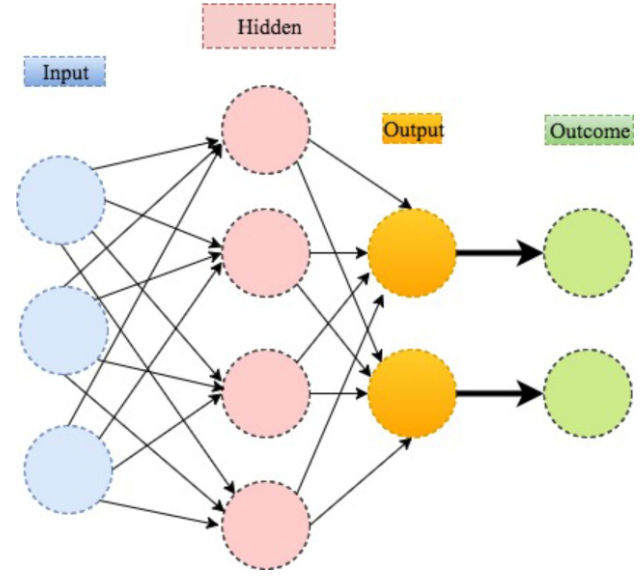
SVM-ի կարևոր հատկությունը կայանում է նրանում, որ մոդելը օպտիմալացվում են ուռուցիկ ֆունկցիայի պարամետրերը, ուսի մենք ձգտում ենք հասնել գլոբալ օպտիմումի։ Մի շարք ուռուցիկ ֆունկցիաներ են օգտագործվում SVM մեթոդում, ուստի այն կիրառելի է շատ բժշկական հետզոտություններում։ SVM-ը կիրառվել է ախտորոշելու նյարդաբանկան և հոգեբուժական հիվանդություններում, քաղցկեղի ախտորոշման համար, Ալցհայմերի ախտորոշման համար և այլն։

**Նեյրոնային ցանցեր**

Նեյրոնային ցանցերը կարելի է ներկայացնել, որպես ռեգրեսիաների ընդլայնված՝ զարգացած տաբերակներ, որոնցով կառուցվում են ոչ գծային կոմբինացիաներ մուտքագրվող փոփոխականի (input) և ելքագրվող (output) փոփոխականի համար միջև։ Նեյրոնային ցանցերում մուտքագրվող և ելքագրվող փոփոխականների միջև կապ է հաստատվում «թաքնված» շերտերի (hidden alyers) կոմբինացիենրի միջոցով։ Նպատակն է փոփոխականների այնպիսի գործակիցների գտնել, որ կանխատեսվոց փոփխականների և իրական արժեքների միջև լինի մինիմալ շեղում։

Նեյրոնային ցանցերը օգտագործվել են ինսուլտի կանխատեսման համար ևս։ Ենթադրենք ներմուծվող փոփոխականները , ինսուլտի սիմպտոմներն են` վերջույթների պարալիզացումը, տեսողության անկումը, խնդիրները շարժունակության հետ և այլն։ Ելքային փոփոխականը՝ ն բինար փոփոխական է։ ցույց է տալիս հիվանդը ինսուլտի է ենթարկվել թե ոչ։ Արդյունքային պարամետրը որոշվում է հետևյալ ֆունկցիայի միջոցով՝

(1.1) հավասարման մեջ փոփոխականների ազդեցության չափերն են, ֆունկցիաների միջոցով տարբեր կոմբինացիաները կանխորոշում են հիվանդության ռիսկը։

**

Գծապատկեր 4. Հասարակ նեյրոնային ցանցի պատկերումը

Մոդելի նպատակն է այնպիսի գտնել, որոնցով կմինիմիզացվի մոդելի «կորստի» ֆունկցիան՝ : Մոդելի կողմից որոշված միջոցով կարելի կանխորոշել միևնույն բնակչության զանգվածին պատկանողհիվանդներին։

Նեյրոնային ցանցեր են կիրառվել կանխատեսելու համար քացկեցը, որտեղ ներմուծվող փոփոխակննեն են եղել  6567 գեները, իսկ արդյունքային փոփոխականը՝ ուռուցքի մակարդակը։ Պակինսոնի հիվանդությունը, որը իրենից ներկայացնում է նյարդային համակարգի խանգարում և շարժողական համակարգի արգելակում, նույնպես ախտորոշվել է նեյրոնային ցանցերի միջոցով։ Այս դեպքում նեմուծվող փոփոխականներ են ծառայել նեյրոնկարները։

Խորը ուսուցումը (Deep learning) դասական նեյրոնային ցանցերի ընդարձակումը կարելի է համարել։ Կարելի համարել խորը ուսուցում երկուսից ավելի «թաքնված» շերտերով նեյրոնային ցանցերով մոդելների ուսուցում։ Ժամանակակից համակարգիչների հզորությունները թույլ են տալիս խորը ուսուցմանը կառուցել բազմաթից շերտերով նեյրոնային ցանցեր, ինչը անհասանելի է դասական նեյրոնային ցանցերի համար։ Այդպիսով խորը ուսուցումը կարող է տվյալների ավելի բարդ ոչ գծային համակցություններ ուսումնասիրել։ Խորը ուսուցման այդքան կիրառելի պատճառներից մեկը դա տվյալների օրեցօր բարդացող և մեծ ծավալների հասնող քանակն է։ Խորը ուսուցումը ամենաշատը կիրառվում է նկարների դասակարգման համար, ուստի բժշկական նկարների ախտորոշման համար այն կայուն գործիք է համարվում։

Բժշկական հետազոտություններում հիմնական կիրառվող նեյրոնային ցանցերն են CNN-ները (convolution neural network) RNN-ները (recurrent neural network) deep belief network-ները և deep neural network-ները։

CNN-ները դասական մեքենայական ուսուցման ալգորիթմների ավելի զարգացած մոդլեներն են, երբ գործ ենք ունենուն մեծ ծավալի հասնող տվյալների հետ։ Դասական մեքենայական ուսուցման ալգորիթմները օգտագործվում են, երբ փոփոխականները ունեն քիչ բնութագրող հատկություններ, իսկ բժշկական հետազոտության համար առկա նկարները շատ մեծ ծավալ ունեն, քանի որ յուրաքանչյուչ նկար իրենից ներկայացնում է հազարավոր փիքսելների հավաքածու, ուստի այս դեպքում հարմար և արդյունավետ է օգտագործել CNN-ներ։

Նկարների վերլուծության ժամանակ (image processing) CNN-ներում ներմուծվող փոփոխականները իրենցից ներկայացնում են նորմալիզացված նկարների փիքսելներ։ Ելքային արդյունքը ստանալու համար CNN-ը փիքսելները վերափոխում է նկարների՝ միջանկյալ շերտերում տարբեր ձևափոխություններ անելով։ Վերջնական արդյունքը ռեկուրսիվ ֆունկցիա է կշռված ներմուծվող փոփոխականներից։ Կշիռները այնպես են «վարժեցվում» (to train), որ կանխատեսվոց փոփոխականների և իրական արժեքների միջև շեղումը հնարավորինս փոքր լինի։

Ինչպես արդեն նշվել է CNN-ները լայն կիրառություն ունեն բժշկության մեջ։ Աչքի նկարների միջոցով հնարավոր է եղել ախտորոշել բնածին կատարախտ հիվանդությունը 90% ճշտությամբ։ Փորձ է արվել նաև ախտորոշել մաշկի քաղցկեղը՝ նկարների միջոցով։ Այս դեպքում ևս ստցվել 90% ճշտություն, և բարորակ ուռուցքների համար, և չարորակ ուռուցքների համար։ Բոլոր այս օրնակները փաստում են, որ CNN-ների դերը բժշկական ախտորոշումների մեջ բավականին մրցակցային է՝ բժիշկների հետ համեմատած։

**Natural language processing**

Նկարները, էլեկտրոֆիզիոլոգիական և գենետիկայի տվյալները, համակարգիչների կողմից հասկանալի տվյալներ են, ուստի նրանցով հնարավոր է կառուցել մեքենայական ուսուցման ալգորիթմներ։ Սակայն բժշկական տվյալների մեծ մասը հաքագրվում է տարբեր տեսակի տեքստերի միջոցով՝ ֆիզիկական հետազոտության արդյունքները, լաբորատոր հետազոտությունների արդյունքերը և այլն։ Natural language processing (NLP)-ի միջոցով հնարավոր է դառնում այս բոլոր տեքստային տվյալներից առանձնացնել այն տվյալները, որոնք հետագայում կաող են ծառայել որպես հիմք որևէ ախտորոշում տալու համար։

NLP-ին կիրառվում է հիմնականում երկու նպատակի համար՝ տեքստի վերամշակման համար և կլասիֆիկացիայի համար։ Տեքստի վերամշակման շնորհիվ NLP-ին կարողանում է հիվանդություններին վերաբերող մի շարք էական բառեր առանձնացնել կլինիկական արձանագրություններից և պատմական բժշկական տվյալներից։ Այնուհետև հավաքված բառերի ամբողջությունը օգտագործվում է էական և ոչ էական դեպքերը դասակարգելու համար։

NLP-ի մեթոդների միջոցով հնարավոր է եղել շուտ ահազանգել անհրաժեշտ բժշկական միջոցառումները իրականցելուն, վերահսկել բուժման աղդեցության հակառակ՝ վնասակար ազդեցությունները և այլն։ Օրինակի համար NLP-ի միջոցով ռենտգենյան զեկույցների մշակումը բժիշկներին ահազանգել է հակաինֆեկցիոն թերապիայի անհրաժեշտությանը։ Հնարավոր է եղել ավտոմատ վերահսկել լաբորատոր ախտորոշման արդյունքների անբարենպաստ լինելը։ NLP-ի ալգորիթմների միջոցով նույնպես հնարավոր է հիվանդություններ ախտորոշել։ ՈՒղեղային անիվրիզմների հիվանդությունը հնարավոր է եղել ախտորոշել NLP-ի միջոցով՝ ստանալով 95% և 86% ճշգրտություններ համապատասխանաբար “training” և “validation” տվյալների բազաների համար։ NLP-ի միջոցով հնարավոր է եղել առանձնացնել երկրորդային զարկերակների հիվանդության մեջ հանդիպող էական բառերը և դասակարգել հիվանդներին, որոնք ունեն այս հիվանդությունը՝ հասնելով 90% ճշգրտության։

Ամփոփելով կարելի է ասել, որ արհեստական բանականության միջոցով հնարավոր է եղել ուսումնասիրել և ախտորոշել մի շարք հիվանդություններ։ Ժամանակակից արհեստական բանականությունը մեքենայական ուսուցման շնորհիվ կարողանում է հետազոտել կառուցվածքային տվյալները, իսկ NLP-ի միջոցով՝ ոչ կառուցվածքայինները։

# **ԳԼՈՒԽ 2. ԱՐՀԵՍՏԱԿԱՆ ԲԱՆԱԿԱՆՈՒԹՅԱՆ ԿԻՐԱՌՈՒԹՅԱՄԲ ՇԱՔԱՐԱՅԻՆ ՌԵՏԻՆՈՊԱԹԻԱՅԻ ԱԽՏՈՐՈՇՈՒՄԸ**

# **2.1 Շաքարային ռետինոպաթիա հիվանդությունը**

Շաքարային ռատինոպաթիան (Diabetic Retinopathy) աստիճանաբար վատթարացող հիվանդություն է, որը հանգեցնում է տեսողության խանգարման և կուրության։ Հիվանդության տարբեր փուլերի միջև աննշան տարբերությունը և փոքր չափերի էական բաղադրիչների առկայությունը շատ են դժվարեցնում հիվանդության ախտորոշումը։ Բացի դրանից հիվանդության ախտորոշումը հիմնականում կատարվում է լաբորատոր պայմաններում, որը ժամանակար պրոցես է և բնականբար ենթադրում է համապատասխան մասնագետի ներգրավվածությունը։ Հիվանդության ախտորոշումը արհեստական բանականության միջոցով էական է այս խնդիրների լուծման համար։ Հիավնդության վաղաժամ ախտորոշումը շատ կարևոր է, քանի որ հայտնաբերելով հիվանդությունը վաղ փուլում՝ հնարավոր է խուսափել կուրությունից։

Շաքարային ռատինոպաթիան այսօր համավում է ամենատարածված հիվանդությունը տեսողության խնդիրներ առաջացնող հիվանդությունների շարքում։ Այս հիվանդությունը առաջանում է շաքարային դիաբետից։ Աչքի ցանցաթաղանթի (retina) արյունատար անոթները ամենաշատն են վնասվում այս հիվանդության ժամանակ։ Հիվանդության սկզբնական փուլում ցանցաթաղանթի անոթները վնասվում են և արյունը լցվում է աչքի մեջ, հաջորդ փուլում հիվանդը հասնում է կուրության։

Ընդհանուր առմամբ շաքարային ռատինոպաթիան կարելի է բաժանել չորս խմբերի՝ կայուն, բավարար, ծանր,ծայրահեղ ծանր։ Առաջին փուլում ցանցաթաղանթի արյունատար անոթները սկսում են ուռչել, երկրորդ փուլում առաջանում է անոթների արյունազեղում, և արյունը այլևս չի մատակարարվում ցանցաթաղանթին։ Երրորդ փուլում վնասված անոթների թիվը շատանում է, ինչն էլ հանգեցնում է վերջին փուլին՝ կուրցմանը։ Ներկայում բժիշկները խորը ուսումնասիրության են ենթարկում աչքի ցանցաթաղանթի նկարները ախտորոշում տալու համար։ Քանի որ ուսումնասիրությունները կատարվում բժիշկների կողմից անհատական ուսումնասիրությունների միջոցով, անհրաժեշտություն է առաջանում ունենալ ավտոմատացված ախտորոշման համակարգը, որը կկարողանա կատարել համընդհանուր հնարավորինս ճիշտ ախտորոշում։

Սկզբնական փուլում ցանցաթաղանթի վնասումը կարող է աննշան լինել, և դժվար բացահայտվող սակայն հետագա բուժման համար շատ կարևոր է հայտնաբերել հիվանդությունը հենց նախնական փուլում։ Ախտորոշումը սկզբնական փուլում դժվար է իրականցնել, քանի որ հարկավոր ուսումնասիրել մի շարք միկրոհատկություններ աչքի նկարների միջոցով, սական հիվանդության փուլերի ճիշտ ախտորոշումը հնարավորություն կտա հնարավորինս խուսափել ավելորդ միջամտությունից, ինչն էլ իր հերթին կկրճատի տեսողության կորցնելու հավանականությունը։

Հիվանդության ավտոմատացված ախտորոշումը սկզբում հիմնված է եղել մեքենայական ուսուցման ստանդարտ ալգորիթմների մոդելների վրա։ Այս պարագայում քանի որ մոդելների համար անհրաժեշտ տվյալները հավաքագրվում էին բժիշկների անհատական ուսումնասիրություններից, ինչը բավական բարդ պրոցես է համարվում, քանի որ անհրաժեշտ է գույնավոր նկարի միկրոմասերին տալ ճիշտ ախտորոշում, այս մեթոդները այնքան էլ ճշգրիտ չէին։ Բացի այդ հնարավոր չէր ճշգրիտ տվյալներ ստանալ, քանի որ բժիշկների տրամադրում տվյալները ոչ որակով նկարներից՝ նկարները կարող էին լինել սխալ անկյան տակ, խավար այլն։ Ուստի այս խնդիրներից խուսափելու համար անհրաժեշտություն առաջացավ ստեղծել այնպիսի ավտոմատացված հակարգը, որը չնայած այս խնդիրների հնարավորինս ճիշտ ախտորոշում կկարողանա տալ։

Վերջին ժամանակներում computer vision-ի մի շարք խնդիրներ լուծվել են խորը ուսուցման ալգորիթմների միջոցով՝ ապահովելով պարձր ճշգրտություն։ CNN-երի միջոցով հնարավոր է եղել տարբեր բնակագավառներում հասնել հաջողությունների, իչպիսիք են առարկայի հայտնաբերումը և հետևումը (object detection and tracking), նկարների և հիվանդությունների դասակարգումը, գործողությունների ճանաչումը (action recognition) և այլն։ CNN-երի հիմնական առավելությունը կայանում է նրանում, որ օբյեկտների հատկությունների դասակարգումը կատավում է ավտոմատ կերպով, ուստի շաքարային ռետինոպաթիայի ախտորոշումը հնարավորինս ճշգրիտ հնարավոր կլինի անել CNN-երի միջոցով։

Հաջորդիվ կներկայացվեն այն մոդելները որոնք օգտագործվել են աշխատանքի շրջանակներում՝ խնդրի լուծման համար։

# **2.2 Հիմնական բաղադրիչի վերլուծության մեթոդ (PCA)**

PCA-ը տվյալների չափողականությունը կրճատելու ալգորիթմ է։ PCA-ը հիմնականում օգտագործվում է