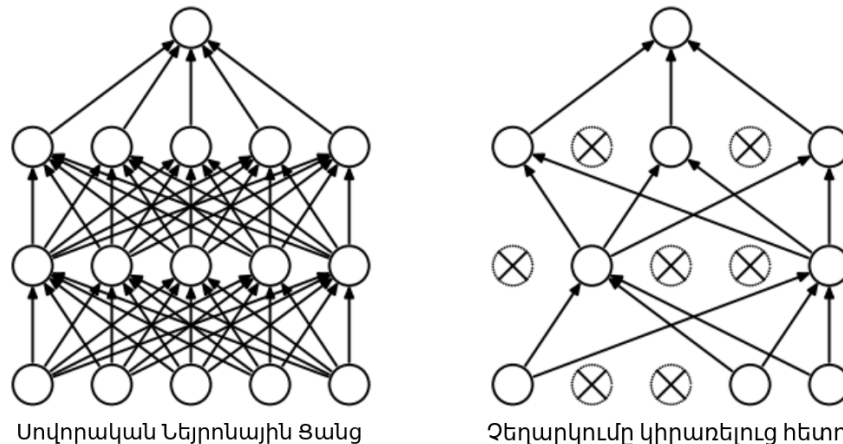


Չեղարկում

Հայկ Կարապետյան

Արդեն ծանոթացել ենք մոդելի գերուսուցված վիճակին և դրա դեմ պայքարելու մեթոդներից ռեգուլարիզացիային: Գերուսուցումից խուսափելու եղանակներից է չեղարկումը (dropout): Նեյրոնային ցանցում ունենք 10 նեյրոնից բաղկացած շերտ: Չեղարկումը առաջարկում է նեյրոնների մի մասը պատահական կերպով անջատել (Նկար 1): Դա կօգնի նեյրոններին շատ չհարմարվել մեր տվյալներին: Այսինքն ամեն անգամ, տարբեր նեյրոնների միջոցով պետք է կայացվի ճիշտ որոշում (նկարի պիտակ): Դրա շնորհիվ կարող ենք խուսափել գերուսուցումից: Նեյրոնների պատահական կերպով անջատումը կատարվում է նրանց կշիռները 0 դարձնելու միջոցով: Չեղարկումը ոչ միայն կարող է օգնել գերուսուցման ժամանակ, այլևս և ինֆորմացիա կորզելու լավ մեթոդ է: Դիտարկենք մեկ օրինակ, որը շատ լավ բնութագրում է չեղարկումը: Պետք է իմանանք ով է սպանել կարմիր կատվին: Ունենք ականատեսների երկու խումբ: Առաջին խմբում կա 8 մարդ, երկրորդ խմբում 15 մարդ: Երկու խմբում էլ կարող ենք անցկացնել 8 մարդուց բաղկացած հարցաքննություն, քանի անգամ ցանկանում ենք: Առաջին խմբում, մի քանի անգամ հարցաքննություն անցկացնելը ցույց տվեց, որ ականատեսների պատասխանները միշտ նույնն են, իսկ երկրորդ խմբում ամեն անգամ պատահական 8 մարդու ընտրելուց և հարցաքննություն անցկացնելուց հետո ունեցանք ավելի տարատեսակ պատասխաններ և կարող ենք ավելի ճշգրիտ որոշում կայացնել: Երկու խմբում էլ հարցաքննում ենք 8 մարդու մի քանի անգամ, այսինքն ծախսում ենք նույն ժամանակը և նույն էներգիան, բայց երկրորդ խմբից ստացանք ավելի շատ ինֆորմացիա: Չեղարկումը ունի մեկ հիպերպարամետր, դա չեղարկման գործակիցն է (drop_rate), որը ցույց է տալիս, թե նեյրոնների քանի տոկոսը պետք է անջատվի:



Նկար 1: Չեղարկումը կիրառելուց հետո նեյրոնների մի մասը անջատվում են

Ուսուցման ժամանակ շերտից դուրս եկող արժեքները բազմապատկվում են $\frac{1}{1-drop_rate}$ -ով: Այսինքն, ակտիվ մնացած նեյրոնների ուժը մեծացնում ենք տալով նրանց այնքան հզորություն, ինչքան կլինեք առանց անջատելու: Հարցաքննության օրինակի դեպքում, եթե ենթադրենք, որ ամեն ականատեսը մեզ թիվ է ասում, երկրորդ խմբում մենք կգումարենք ստացած 8 թվերը և կբազմապատկենք $\frac{15}{8}$ -ով: Մոդելի փորձարկման (inference) ժամանակ չեղարկումը չի կարող օգտագործվել, քանի որ մոդելում չեն կարող պատահական արժեքներ մնալ: Բոլոր նեյրոնները միացվում են, չեղարկման գործակիցը հավասար է 0-ի և արդյունքում պատահական ոչինչ չի մնում: