



Դավիթ Սպարտակի Քարամյանի
«խոսքի հուսալի մշակում՝ ներդրումն անհետնական
բանականության կիրառություններում» վերնագրով:

Ե.13.05- «Մաթեմատիկական մոդելավորում, թվային մեթոդներ և օրոգրների համալիրներ»
մասնագիտությամբ:

տեխնիկական գիտությունների թեկնածուի
գիտական աստիճանի հայցման ատենախոսության մասին

ԱՌԱՋԱՏԱՐ ԿԱԶՄԱԿԵՐՊՈՒԹՅԱՆ ԿԱՐԾԻՔ

Ատենախոսությունը նվիրված է անգլերեն խոսքի ճանաչման և խոսնակների տարբերակման (դիարիզացիայի) խնդրի համար մեքենայական ուսուցման մեթոդներով աշխատող արագագործ համակարգի մշակմանը: Համակարգը բաղկացած է մեծ թվով բաղադրիչներից, որոնցից վեցում առաջարկվել են որոշակի լավարկումներ, ինչպես որակի, այնպես էլ արագագործության տեսանկյունից:

Ատենախոսությունը բաղկացած է չորս գլխից, ներառյալ ներածությունը, եզրակացությունից, հապավումների ցանկից և գրականության ցանկից:

Առաջին գլխում որպես ներածություն նկարագրված է խոսքի ճանաչման և խոսնակների տարբերակման խնդիրների բարդությունները, ամփոփված է աշխատանքի նպատակը և հիմնական նորոյթները:

Երկրորդ գլխում լավագույն ձևով ներկայացված է խնդրի նախապատմությունը, անհրաժեշտ գաղափարները, կիրառվող հիմնական մոտեցումները ու ալգորիթմները: Մասնավորապես, համակարգված ձևով ներկայացված են Mel հատկանիշները, դրանց ուժեղ և թույլ կողմերը, նեյրոնային ցանցերում կիրառվող կորստի ֆունկցիաները, դեկոդավորման մոտեցումները, խոսնակների նույնականացման ժամանակակից մեթոդները, խոսքի ճանաչման և խոսնակների տարբերակման ալգորիթմների գնահատման մեթոդները:

Երրորդ գլխում ներկայացված են ալգորիթմական հիմնական նորոյթները, որոնք ուղղված են խոսքի ճանաչման և խոսնակների տարբերակման ալգորիթմի որակի բարձրացմանը տարբեր առանքներով: 3.1 ենթագլխում ներկայացված է ձայնագրության մեջ բնական աղմուկների ավելացման մեթոդները: 3.2 ենթագլխում նկարագրված է խոսքի ճանաչման հիմնական մեթոդի ուսուցումը, որը հիմնված է հայտնի Conformer ճարտարապետության վրա: Մասնավորապես տպավորիչ է 3.1 աղյուսակը, որտեղ ցույց է տրված, որ ստացված շատ չի զիջում Meta և OpenAI ընկերությունների ստեղծած մոդելներից անգլերեն խոսքի ճանաչման OpenASR շտեմարանի վրա, չնայած որ էապես ավելի փոքր և արագագործ մոդել է:

3.3 ենթագլխում նկարագրված է բնական աղմուկի պայմաններում խոսքի ճանաչման ալգորիթմի լավացման մի մեխանիզմ, որի հիմքում աղմուկի մասնակի հեռացումն է: Մի շարք

Էքսպերիմենտներով ցույց է տրված, որ արդեն նախապես ուսուցված մոդելի կայունությունը բնական աղմուկների նկատմամբ կարելի է լավագույն բարձրացնել կատարելով աղմուկի միայն մասնակի հեռացում: 3.4 ենթագլխում դիտարկված է խոսքի ճանաչման ալգորիթմի ադապտացումը օգտագործողի կողմից տրվող բառարանի բառերին՝ առանց համակարգը նորից ուսուցանելու: Դիտարկված են մի շարք մոտեցումներ և կատարված են մեծ թվով փորձեր: Արդյունքում առաջարկվել է ալգորիթմ, որը գերազանցում է OpenAI Whisper Small-ին և համեմատելի է OpenAI Whisper Large ալգորիթմի հետ, լինելով էապես ավելի արագագործ:

3.5 ենթագլխում դիտարկված է խոսնակների նույնականացման մոդելի՝ բնական աղմուկների նկատմամբ կայունության ավելացման խնդիրը: Առաջարկվել են մի շարք մոտեցումներ, որոնք էապես բարելավել են համակարգի կայունությունը: Մասնավորապես հետաքրքիր է teacher-student սխեմայով աշխատող ալգորիթմը: Վերջապես, 3.6 ենթագլխում ուսումնասիրվել է խոսնակների տարբերակման ալգորիթմի սխալները ուղղելու մի քանի մոտեցումներ: Ցույց է տրվել, որ առաջարկված մեթոդներից մեկը բառերի դասակարգման սխալանքը բարելավում է 0.12 տոկոսային կետով:

Չորրորդ գլուխը ներկայացված են համակարգում արված չորս փոփոխություններ, որոնք էապես մեծացնում են համակարգի արագագործությունը: Մուտքային տվյալների չափերի կրճատումը, նեյրոնային ցանցի քվանտիզացիան, լեզվական մոդելի փոխարինումը պարզագույն N-gram մոդելով և կլաստերինգի ալգորիթմի փոխարինումը երկբայանի ալգորիթմով էապես չեն թուլացնում համակարգի որակը, սակայն այն դարձնում են կիրառելի առօրյայում օգտագործվող ցանկացած համակարգչի համար:

Այսպիսով, Դավիթ Զարամյանի ատենախոսության մեջ առաջարկված են խոսքի ճանաչման և խոսնակների նույնականացման ամբողջական համակարգի գրեթե բոլոր բաղադրիչների լավացումներ: Աշխատանքը շարադրված է բարձր գիտական մակարդակով: Հիմնական արդյունքները տպագրված են գիտական հոդվածներում: Սեղմագիրը համապատասխանում է ատենախոսության բովանդակությանը:

Աշխատանքում նկատված թերացումները.

- Աշխատանքում նշվում է, որ նման բազմաբաղադրիչ համակարգերում տեղի է ունենում սխալների կուտակում, նաև հակիրճ անդրադարձ կա end-to-end լուծումներին, որոնք առայժմ արդյունավետությամբ զիջում են նկարագրված համակարգին: Աշխատանքը կշահեր, եթե որոշակի ջանք դրվեր այդպիսի համակարգերի ուսումնասիրության ու լավարկման փորձերի վրա, քանի որ մեծ հավանականությամբ դրանք են լինելու ապագայի լավագույն մոդելները:
- 3.5 ենթագլխում խոսնակների վեկտորների որոշման ալգորիթմում կիրառվել է teacher-student ալգորիթմ, որտեղ teacher-ի պարամետրերը ուսուցման երկրորդ փուլում չեն փոփոխվում: Նմանատիպ ճարտարապետությունները շատ տարածված են պատկերների ճանաչման ոլորտում և ունեն զգալի բարելավումներ: Մասնավորապես, Mean Teacher, BYOL, DINO ալգորիթմներում teacher ցանցի պարամետրերը պարբերաբար թարմացվում են student ցանցի պարամետրերից exponential moving average սխեմայով: Այսպիսի մոտեցումների ադապտացումը խոսնակների ճանաչման ցանցի համար կարող է օգտակար լինել:
- 4.4 ենթագլխում խոսնակների կլաստերիզացիայի համար հետաքրքիր կլինեն հասկանալի, թե ինչ արագությամբ են աշխատում նոսր մատրիցների eigendecomposition ալգորիթմները, օրինակ scipy գրադարանի `scipy.sparse.linalg.eigs` ֆունկցիան:

Այս դիտողությունները, սակայն, չեն ազդում աշխատանքի գիտական արժեքի վրա:

Կարծում ենք, որ ատենախոսությունը բավարարում է ՀՀ ԲԿԳԿ-ի կողմից Ե.13.05- «Մաթեմատիկական մոդելավորում, թվային մեթոդներ և ծրագրերի համալիրներ» մասնագիտությամբ թեկնածուական ատենախոսություններին ներկայացվող պահանջներին, և նրա հեղինակը՝ Դավիթ Զարամյանը, արժանի է տեխնիկական գիտությունների թեկնածուի գիտական աստիճանի շնորհմանը:

Աշխատանքը բննարկվել է ԵՊՀ Դիսկրետ մաթեմատիկայի և տեսական ինֆորմատիկայի ամբիոնի սեմինարին: Զննարկմանը ներկա էին ԻԿՄ ֆակուլտետի դեկան, ֆ.մ.գ.դ. պրոֆեսոր Յու. Հակոբյանը, Դիսկրետ մաթեմատիկայի և տեսական ինֆորմատիկայի ամբիոնի վարիչ, ֆ.մ.գ.թ. դոցենտ Պ. Պետրոսյանը, ֆ.մ.գ.դ. Ա. Զուբարյանը, ֆ.մ.գ.թ. Ս. Սարգսյանը, տ.գ.թ. Ա. Հովակիմյանը, դոցենտ Է. Դանոյանը, ֆ.մ.գ.թ. Հ. Խաչատրյանը, ֆ.մ.գ.թ. Վ. Հուրոյանը:

ԵՊՀ Դիսկրետ մաթեմատիկայի և տեսական
ինֆորմատիկայի ամբիոնի վարիչ,
ֆ.մ.գ.թ. դոցենտ

Պ. Պետրոսյան

ԵՊՀ Ինֆորմատիկայի և կիրառական
մաթեմատիկայի ֆակուլտետի դեկան,
ֆ.մ.գ.դ. պրոֆեսոր

Յու. Հակոբյան

սպորտաբարձր Սեր Խաչատրյան ԺԺ



Ք. Բաբայան