

**ԽՈՍՔԻ ԱՎՏՈՄԱՏԱՑՎԱԾ ՃԱՆԱՉՄԱՆ
ԿԻՐԱՌՈՒԹՅՈՒՆՆԵՐԸ ՏՆՏԵՍՈՒԹՅՈՒՆՈՒՄ
ԵՎ ՀԱՅԵՐԵՆ ԽՈՍՔԻ ՁԱՅՆԱՅԻՆ ՏՎՅԱԼՆԵՐԻ
ԲԱԶԱՅԻ ՍՏԵՂԾՄԱՆ ԱՆՀՐԱԺԵՇՏՈՒԹՅՈՒՆԸ**

ՏԻԳՐԱՆ ՔԱՐԱՄՅԱՆ

Բանալի բառեր. բնական լեզվի մշակում, խոսքի ավտոմատացված ճանաչում, խոսքի անոտացիա, «Պատմվածքներ-15» տվյալների բազա, նմուշառման արագություն, ոչ կառուցվածքային տվյալներ:

ՆԵՐԱԾՈՒԹՅՈՒՆ

Խոսքի ճանաչումը, որը նաև հայտնի է որպես խոսքի ավտոմատացված ճանաչում (Automated Speech Recognition) թույլ է տալիս մարդկային բանավոր խոսքը վերածել գրավորի: Այն համարվում է համակարգչային գիտության ամենաբարդ ոլորտներից մեկը՝ ներառելով լեզվաբանությունը, մաթեմատիկան և վիճակագրությունը: Խոսքի ճանաչման մոդելները կազմված են մի քանի բաղադրիչներից, ինչպիսիք են խոսքի մուտքագրումը, հատկանիշի դուրս բերումը (feature extraction), հատկանիշի վեկտորները, ապակոդավորիչը (decoder) և էլքային բառը, որտեղ ապակոդավորիչը օգտագործում է ակուստիկ մոդելներ, արտասանության բառարաններ և լեզվի մոդելներ՝ համապատասխան էլքային արդյունքը գտնելու համար: Մոդելների ճշգրտության գնահատման հիմնական ցուցանիշներն են բառի սխալի մակարդակը (Word Error Rate) և մոդելի արագությունը: Մակայն, մի շարք գործոններ կարող են ազդել բառերի սխալի մակարդակի վրա, ինչպիսիք են արտասանությունը, շեշտը, ձայնի բարձրությունը և ֆոնային աղմուկը: Խոսքի ճանաչման բազմաթիվ հավելվածներ և ծրագրեր կան, բայց ավելի առաջադեմ լուծումներն օգտագործում են արհեստական բանականության (AI) և մեքենայական ուսուցուման (ML) գործիքակազմը: Արհեստական բանականությունը և մեքենայական ուսուցումը միավորում են ձայնային ազդանշանների քերականությունը, շարահյուսությունը, կառուցվածքը և կազմը՝ մարդկային խոսքը հասկանալու և մշակելու համար: Ներկայումս խոսքի ճանաչման ամենատարածված ալգորիթմներն ու մոտեցումներն են.

- Բնական լեզվի մշակում (NLP). NLP-ն արհեստական բանականության ոլորտ է, որը կենտրոնանում է մարդկանց և մեքենաների փոխազդեցության վրա խոսքի և տեքստի միջոցով: Շատ շարժական սար-

քեր, օրինակ սմարթֆոններ, իրենց համակարգերում ներառում են խոսքի ճանաչում՝ որոնք թույլ են տալիս իրականացնել որոշակի գործողություններ ձայնի միջոցով, օրինակ. Siri կամ Bixby:

- N-գրամներ¹. N-գրամները լեզվի պարզագույն մոդել է, որը հավանականություններ է վերագրում նախադասություններին կամ արտահայտություններին: N-գրամը N-բառերի հաջորդականությունն է և այդ մոդելը օգտագործվում է տեքստի քերականությունը և որոշակի բառային հաջորդականությունների հավանականությունը հասկանալու համար, ինչը խոսքի ճանաչումը դարձնում է ավելի ճշգրիտ:

- Թաքնված Մարկովյան մոդելներ (Hidden Markov Models)². Խոսքի ճանաչման ժամանակակից ընդհանուր նշանակության համակարգերը հիմնված են թաքնված Մարկովյան մոդելների վրա, որոնք պատկանում են վիճակագրական մոդելների դասին: Մարկովյան մոդելները հիմնված են Մարկովի շղթայի մոդելի վրա, որը սահմանում է, որ տվյալ վիճակի հավանականությունը կախված է ներկա վիճակից, այլ ոչ թե նրա նախորդ վիճակներից, իսկ թաքնված Մարկովյան մոդելները թույլ են տալիս ներառել թաքնված իրադարձությունները, ինչպիսիք են խոսքի մասի պիտակները (part of speech), հավանական մոդելի մեջ: HMM-ները օգտագործվում են խոսքի ճանաչման մեջ, քանի որ խոսքի ազդանշանը կարող է դիտարկվել որպես մասնակի ստացիոնար ազդանշան կամ կարճաժամկետ ստացիոնար ազդանշան: Կարճ ժամանակային ինտերվալների պարագայում (օրինակ՝ 10 միլիվայրկյան), խոսքը կարող է համարվել ստացիոնար գործընթաց, հետևաբար ամբողջական խոսքը կարելի է դիտարկել որպես Մարկովյան մոդել բազմաթիվ ստոխաստիկ գործընթացների համար:

- Նեյրոնային ցանցեր. նեյրոնային ցանցերը սովորում են վարժեցվող տվյալների (training data) հիման վրա՝ ընդօրինակելով մարդկային ուղեղի նյարդային հանգույցների փոխկապակցվածությունը նեյրոնային շերտերի միջոցով: Յուրաքանչյուր հանգույց կազմված է մուտքերից, կշիռներից, շեղումից (bias) և ելքից: Հանգույցը «գործարկվում» կամ ակտիվացվում է կախված տվյալ հանգույցի ելքային արժեքից և փոխանցում է տվյալները ցանցի հաջորդ շերտին: Թեև նեյրոնային ցանցերը հակված են ավելի ճշգրիտ լինել և կարող են ավելի շատ տվյալներ ընդունել և մշակել, սակայն նմանատիպ մոդելների կառուցումը բավականին ռեսուրսատար է՝ համեմատած ավանդական լեզվի մոդելների հետ:

¹ Jurafsky, D., Martin H.M. Speech and Language Processing: An Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics (3rd edition). – 2019. – P. 30-56, Link.

² Aggarwal, C. C. Machine learning for text Cham: Springer. -2018. -Vol. 848. -P. 387-396, Link.

• End-to-end¹ խոսքի ճանաչման մոդելներ. ավանդական հնչյունական (օրինակ՝ HMM-ի վրա հիմնված բոլոր մոդելները) մոտեցումները պահանջում էին առանձին բաղադրիչների ուսուցում արտասանության, ակուստիկ և լեզվի մոդելի համար: End-to-end մոդելները համատեղ սովորում են խոսքի ճանաչման բոլոր բաղադրիչները, ինչը հեշտացնում է ուսուցման գործընթացը և մոդելի տեղակայման (deployment) գործընթացը:

ՀԱՅԵՐԵՆ ԼԵԶՎԻ ՃԱՆԱՉՄԱՆ ՀԻՄՆԱԽՆԴԻՐՆԵՐԸ

Աշխարհի տարբեր լեզուների հիմնական մասը համարվում է քիչ ուսուցանվող կամ քիչ տարածված: Այդպիսի լեզուներից է հայերենը, որը հիմնականում օգտագործվում է պատմական Հայաստանի տարածքներում և հայկական սփյուռքում: Քիչ ուսուցանվող լեզուների համար խոսքի ճանաչման մոդելներ կամ գործիքներ ստեղծելուց մեծ քանակությամբ ռեսուրսներ (տվյալներ) են հարկավոր, ինչը, սակայն, այդ լեզուների համար ծախսատար կամ դժվարամատչելի է:

Խնդիրներից մեկը բավարար քանակությամբ խոսնակների ձայնային տվյալների հավաքագրումն է, որը անհրաժեշտ է մոդելին՝ ձայնային բոլոր տոնայնությունները հասկանալու համար, քանի որ մարդկանց ձայները միմյանցից տարբերվում են տոնայնությամբ, արտասանության արագությամբ, ձայնային տեմբրով և այլն: Հետևաբար, խոսքի ճանաչման ճշգրիտ մոդել ստեղծելու համար անհրաժեշտ է ճանաչման մոդելը վարժեցնել տվյալների մեծ բազայի վրա, որը իր հերթին պետք է բաղկացած լինի բազմաթիվ խոսնակներից: Ոչ հիմնական լեզուների դեպքում բավականին դժվար է ապահովել այդպիսի տվյալների հավաքագրումը և տեսականորեն անհնար է որոշել, թե որն է խոսնակների բավարար քանակությունը ճշգրիտ մոդելի վարժեցման համար:

Երկրորդ խնդիրը ձայնային տվյալների հնչյունաբանական բազմազանությունն է: Խոսքի ժամանակակից ճանաչման մոդելներից շատերը օգտագործում են վիճակագրական մոդելներ, որոնք ներկայացնում են հնչյունները համատեքստում կամ տրիֆոնների միջոցով, այսինքն՝ մոդելը սովորաբար ներկայացնում է տվյալ հնչյունը որոշակի ձախ հարևան հնչյունի և որոշակի աջ հարևան հնչյունի հետ միասին: Լեզվաբանության

¹ Graves A., Jaitly N. Towards end-to-end speech recognition with recurrent neural networks // In Proceedings of the 31st International Conference on International Conference on Machine Learning. – 2014. – Vol. 32 (ICML'14). JMLR.org. – P. 1764–1772.

² Precoda, K. Non-Mainstream Languages and Speech Recognition: Some Challenges // *the CALICO Journal*. – 2004. – Vol 21, No. 2. – P. 229-243.

մեջ տրիֆոնը երեք իրար հաջորդող հնչյունների հաջորդականությունն է, որոնք օգտագործվում են բնական լեզվի մշակման մոդելներում տարբեր համատեքստերում տվյալ հնչյունը հասկանալու համար: Նույնիսկ ամենափոքր հնչյունային ֆոնդով լեզուներում վավերացված տրիֆոնների թիվը կարող է շատ մեծ լինել: Բացի այդ, բազմաթիվ տրիֆոնների հավաքագրումը պետք է իրականացվի տարբեր համատեքստերից, քանի որ տրիֆոնները չեն ընդգրկում հնչյունական իրացման վրա ազդող բոլոր համատեքստային գործոնները, և մոդելները պետք է ներառեն նաև ալոֆոններ, որոնք պայմանավորված են տրիֆոններից դուրս համատեքստով: Հնչյունաբանության մեջ ալոֆոնը մի շարք հնարավոր խոսակցական ձայների, նշանների այն ամբողջությունն է, որն օգտագործվում է միավոր հնչյունն արտասանելու համար: Մեծաքանակ հնչյունային միավորների պարագայում, զուգորդված խոսնակների զանազանության հետ, հանգեցնում են հազարավոր բառապաշարային արտահայտությունների անհրաժեշտության, որոնք դժվարացնում են տվյալների հավաքագրման գործընթացը:

Երրորդ մարտահրավերը, որը հաճախ հանդիպում է քիչ տարածված լեզուների հետ աշխատելիս էլեկտրոնային բառացանկերի (համապատասխան հնչյունաբանական տառադարձումներով) բացակայությունն է: Խնդիրն այն է, որ խոսքի ճանաչման մոդելների ստեղծման ժամանակ բացի ձայնային տվյալներից անհրաժեշտ են նաև այդ ձայնային տվյալների տեքստային տարբերակները: Լայն տարածում ունեցող լեզուներում նման տեքստային տվյալների հավաքագրումը կարող է հասանելի լինել տարատեսակ էլեկտրոնային բառարաններից, տեքստային տվյալներ տրամադրող հարթակներից և այլն, այնինչ քիչ տարածված լեզուներում նման տվյալների բազաների ստեղծումը լրացուցիչ ռեսուրսներ է պահանջում:

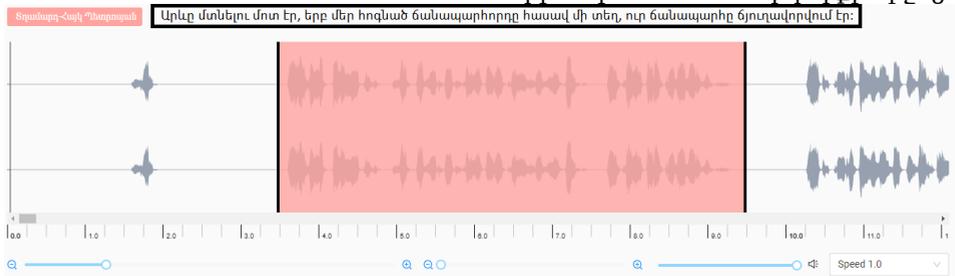
Եվ վերջապես, չորրորդ հիմնական խնդիրը լեզվի ուղղագրական և գրագիտական առանձնահատկություններն են: Որոշ լեզուներում, ինչպես օրինակ հայերենում, մեկ կամ մի քանի լեզվային համակարգեր կարող են գոյություն ունենալ (բարբառներ, գրաբար, արևմտահայերեն, արևելահայերեն և այլն): Բացի այդ ձայնային և տեքստային տվյալները ստանդարտացված չեն, ինչպես, օրինակ, տարբեր գրողների միջև, այնպես էլ նույն գրողի տարբեր աշխատություններում: Այլ կերպ ասած, նույն տառը կամ բառը կարող են տարբեր կերպ արտասանվել տարբեր խոսնակների կողմից: Ավելին, նույն անձը կարող է նույն արտահայտությունը տարբեր կերպ արտասանել կախված տրամադրությունից, համատեքստից, օրվա ժամից և այլն: Ոչ ստանդարտացված ձայնային և տեքստային տվյալների առկայությունը, ինչպես նաև ուղղագրական և արտասանական անճշտությունները կարող են զգալիորեն մեծացնել ձայնային տվյալների հավա-

քագրման և նախապատրաստման աշխատանքների ծավալները, որոնք պետք է օգտագործվեն խոսքի ճանաչման համակարգի ուսուցման համար:

«ՊԱՏՄՎԱԾՔՆԵՐ-15» ՏՎՅԱԼՆԵՐԻ ԲԱԶԱ

«Պատմվածքներ-15» տվյալների բազան¹ իրենից ներկայացնում է հայերեն շուրջ 40 պատմվածքներից հավաքագրված ձայնային և տեքստային տվյալներ և այս բազայի ստեղծման նպատակն է ունենալ հայերեն խոսքի ճանաչման կորպուս, որը թույլ կտա ավելի ճշգրիտ ավտոմատացված խոսքի ճանաչման մոդելներ վարժեցնել: Տվյալների բազայում տեղ գտած բոլոր պատմվածքները ընթերցվում են արևելահայերենով, չնայած որ առանձին պատմվածքներում կան նաև բարբառային արտահայտություններ: Տվյալների բազայի անվանման «15»-ը ներկայացնում է անոտացված (պիտակավորված) ժամերի քանակը: Այստեղ անոտացիան² կամ պիտակավորումը վերաբերում է ձայնային որոշակի միջակայքին համապատասխան տեքստի վերագրումը, որի օրինակը ներկայացված է Գծապատկեր 1-ում: Այլ կերպ ասած, տվյալների բազան բաղկացած է 40 պատմվածքից, որոնք անոտացվել են, և անոտացված ժամանակահատվածը կազմել է մոտավորապես 15 ժամ: 40 պատմվածքների մեջ տեղ են գտել ինչպես արտասահմանյան գրականության թարգմանություններ, այնպես էլ հայ անվանի գրողների տարբեր աշխատություններ:

Գծապատկեր 1. Ձայնային տվյալների անոտացիայի օրինակ՝ Label Studio գործիքի միջոցով



Անոտացիաների երկարությունը տատանվում է 1 վայրկյանից առավելագույնը մինչև 20 վայրկյան: Տարբեր երկարության անոտացիաները թույլ են տալիս ավելի ճշգրիտ հասկանալ հնչյունները, քանի որ նույն բառը (նույն հնչյունը) հանդես է գալիս տարբեր համատեքստերում:

¹ Բոլոր տվյալները հավաքագրվել են grqser.org կայքից:

² Ձայնային տվյալների անոտացիան իրականացվել է Label Studio գործիքի միջոցով:

«Պատմվածքներ-15» տվյալների բազայում տեղ են գտել 12 խոսնակների ձայնագրություններ, որոնցից 9-ը տղամարդկանց ձայնագրություններ են (մոտավորապես 9 ժամ 20 րոպե), իսկ 3-ը՝ կանանց (մոտավորապես 5 ժամ 45 րոպե): Ձայնային տվյալների նմուշառման հաճախականությունը¹ 44100 և ավել է, զերծ տարատեսակ ֆոնային աղմուկներից: Համապատասխան տեքստային ֆայլերը կետադրված են, և պահպանված են հայերենի բոլոր ուղղագրական կանոնները: Տվյալների բազայի համառոտ նկարագիրը ներկայացված է Աղյուսակ 1-ում:

Աղյուսակ 1. «Պատմվածքներ-15» տվյալների բազայի ամփոփ նկարագիր

	Մեռ	Խոսնակ	Խոսքի տևողություն	
	Տղամարդ	Գրիգոր Բաղդասարյան	5 ժամ	55 րոպե
	Տղամարդ	Արթուր Գեղամյան	2 ժամ	2 րոպե
	Տղամարդ	Նարեկ Բաղդասարյան	0 ժամ	19 րոպե
	Տղամարդ	Դավիթ Թորոսյան	0 ժամ	15 րոպե
	Տղամարդ	Հայկ Պետրոսյան	0 ժամ	14 րոպե
	Տղամարդ	Էլֆիկ Զոհրաբյան	0 ժամ	11 րոպե
	Տղամարդ	Անդրանիկ Անտոնյան	0 ժամ	9 րոպե
	Տղամարդ	Նարեկ Բավեյան	0 ժամ	8 րոպե
	Տղամարդ	Վահագն Գրիգորյան	0 ժամ	5 րոպե
0	Կին	Անահիտ Կիրակոսյան	4 ժամ	32 րոպե
1	Կին	Լիդիա Խաչատրյան	0 ժամ	40 րոպե
2	Կին	Կարինե Հովհաննիսյան	0 ժամ	33 րոպե
	Ընդհանուր խոսքի տևողություն		15 ժամ	4 րոպե
Նմուշառման արագությունը (Sampling rate)		Ձայնային տվյալների 97% - 44100 Hz Ձայնային տվյալների 3% - 48000 Hz		

¹ Նմուշառման արագությունը կամ նմուշառման հաճախականությունը սահմանում է մեկ վայրկյանում (կամ միավոր ժամանակահատվածում) նմուշների քանակը, որոնք վերցվում են շարունակական ազդանշանից՝ դիսկրետ կամ թվային ազդանշան ստանալու համար:

ԽՈՍՔԻ ՃԱՆԱՉՄԱՆ ՄՈՂԵԼՆԵՐԻ ԿԻՐԱՌՈՒԹՅՈՒՆԸ

Բազմաթիվ ընկերություններ այսօր օգտագործում են խոսքի ճանաչման տեխնոլոգիաները՝ օգնելով ձեռնարկություններին և սպառողներին խնայել ոչ միայն ժամանակ, այլ նաև ռեսուրսներ: Խոսքի ճանաչման մոդելների ամենալայն կիրառություններից է ձայնային որոնումը. շնորհիվ այնպիսի հավելվածների, ինչպիսիք են Siri կամ Google ձայնային օգնականը, ձայնային որոնումները բավականին լայն տարածում են ստացել: Բավականին տարածված են նաև ձայնից տեքստի վերածելու կիրառությունները և ձայնային հրամաններով տարբեր համակարգերի կառավարումը: Առաջինի դեպքում խոսքի ճանաչումը հնարավորություն է տալիս ձայնի միջոցով մուտքագրելու էլեկտրոնային նամակներ, հաշվետվություններ և այլ փաստաթղթեր, իսկ երկրորդի դեպքում՝ ձայնային հրամանների միջոցով կառավարել մեքենաներ, բնակարան, ռոբոտներ և այլն:

Ֆինանսներ. ֆինանսական բորսաներում այս կամ այն ապրանքի, ծառայության կամ արժեթղթի գների տատանումների վրա մեծ ազդեցություն է ունենում ոլորտի մասնագետների կարծիքը: Հաճախ հարցազրույցների կամ այլ բանավոր փոխհարաբերությունների արդյունքում մասնագետների կողմից արտահայտված կարծիքը որոշակի ազդեցություն է թողնում շուկայում, ինչի պատճառով անհրաժեշտություն է առաջանում տեքստային ֆորմատի վերածել մասնագետների բանավոր խոսքը: Խոսքի ավտոմատացված ճանաչումը թույլ է տալիս լուծել այդ խնդիրը: Ստացված տեքստային տվյալները, հետագա մշակում անցնելուց հետո, կարող են ֆինանսական և տնտեսական մոդելների համար տվյալների բազա ծառայել, այս կամ այն երևույթների միջև ազդեցությունը գնահատելու համար:

Սպասարկման ոլորտ. խոսքի ճանաչման տեխնոլոգիաները տարբեր կիրառություն կարող են ունենալ սպասարկման ոլորտում: Այն կարող է օգնել զանգերի կենտրոնին վերլուծել հազարավոր հեռախոսագանգեր բացահայտելու ընդհանուր զանգերի նպատակներն ու խնդիրները: Բացի այդ, արհեստական բանականության չաթ-բոտերը կարող են նաև զրուցել մարդկանց հետ՝ պատասխանելով հիմնական հարցումներին՝ առանց սպասելու կոնտակտային կենտրոնի գործակալի հասանելիությանը:

Առողջապահություն. առողջապահության ոլորտում խոսքի ճանաչումը կարող է իրականացվել բժշկական փաստաթղթավորման գործընթացում: Հիվանդների զննումների ժամանակ բժիշկները խոսքը տեքստ դարձնող խոսքի ճանաչման մոդելների միջոցով կարող են հիվանդների ախտանշանները գրի առնելու փոխարեն ուղղակի խոսքի միջոցով լրացնել բոլոր փաստաթղթերը, ինչը թույլ կտա ժամանակ խնայել և ավելի

շատ հիվանդների սպասարկել: Որոշ արհեստական բանականության մոդելներ թույլ են տալիս նաև ախտորոշել հիվանդի հոգեկան վիճակը՝ հիվանդի ձայնը վերլուծելով:

Կրթություն. ձայնի ճանաչման տեխնոլոգիաները էական դեր են խաղում նաև կրթության մեջ: Այս տեխնոլոգիան թույլ է տալիս հաշմանդամներին, տեսողական խնդիրներ ունեցող մարդկանց ձայնի միջոցով տեքստային փաստաթղթեր կազմել, ինչպես նաև ձայնի միջոցով կառավարել տարատեսակ առարկաներ, օրինակ համակարգիչը: Ներկայումս բազմաթիվ ծրագրային ապահովումներ կան, որոնք թույլ են տալիս նույնիսկ ձայնային հրամանների միջոցով իրականացնել ֆինանսական վերլուծություններ:

Անվտանգություն. խոսքի ճանաչման մոդելները բարելավում են վարորդների անվտանգությունը՝ թույլ տալով ձայնային հրամանների միջոցով կառավարել նավիգացիոն համակարգերը և մեքենայի աուդիոհամակարգերում առկա որոնման հարթակները: Բացի այդ, ձայնի վրա հիմնված նույնականացումը ավելացնում է տվյալների պահպանման անվտանգությունը: Պետության տեսանկյունից ռազմավարական նշանակություն ունի հարևան երկրների լեզուների ավտոմատացված ճանաչումը և այդ խոսքի վերածումը տեքստի: Ռազմավարական կարևոր նշանակություն ունի «դիմացինին» արագ հասկանալու ունակությունը, ինչը մեծ առավելություն կտա տեղեկատվական պատերազմներում, դիվանագիտության մեջ, ինչպես նաև երկրի պաշտպանության հարցում:

ԵԶՐԱԿԱՑՈՒԹՅՈՒՆ

Խոսքի ճանաչման մոդելների կիրառությունը ավելի ու ավելի լայն տարածում է ստանում, ինչը ոչ միայն հեշտացնում է մարդկանց կյանքը, այլև դարձնում է ավելի ապահով: Բնական լեզվի մշակման մեթոդները առանցքային դեր են խաղում ոչ կառուցվածքային տվյալների հետ առնչվելիս, ինչպիսիք են ձայնային և տեքստային տվյալները: Աշխատանքում ներկայացված են հայերեն խոսքի ավտոմատացված ճանաչման հիմնական մարտահրավերները, որոնցից ամենակարևորը ուսուցման տվյալների բազայի սակավությունն է: «Պատմվածքներ-15» տվյալների բազայի ստեղծումը նպատակ ունի լուծելու տվյալների սակավության խնդիրը և զարգացնելու խոսքի ճանաչման մոդելների ուսուցումը հայերենի համար: Այդ մոդելների կիրառությունների բազմազանությունը թույլ է տալիս դրանք դիտարկել նաև որպես ռազմավարական կարևոր ներդրումներ տնտեսության տարբեր ոլորտներում:

Тигран Карамян, Применения автоматического распознавания речи в экономике и необходимость создания базы данных для армянской речи – Все более широкое распространение получает использование моделей распознавания речи, которые не только облегчают жизнь людей, но и делают ее более безопасной. Методы обработки естественного языка играют ключевую роль при работе с неструктурированными данными, такими как аудио/текстовые данные. Стоит отметить, что методы обработки естественного языка могут выявить, связаны ли чувства, информация, тон или темы с определенными результатами. Эти подсказки можно использовать для построения проверяемых гипотез о том, какие факторы влияют на те или иные решения и как. С другой стороны, эти методы и их применение в разных отраслях позволяют упростить сбор, обработку, передачу и использование информации.

В работе представлены основные проблемы и трудности создания автоматизированных моделей распознавания речи, особенно для менее распространенных языков, а также применения моделей распознавания речи в различных отраслях экономики, таких как финансы, сектор обслуживания, здравоохранение, образование, кибер безопасность. Цель – представить общественности базу данных «Истории-15», которая представляет из себя 15-часовые аудио/текстовые данные, собранные примерно из 40 рассказов на армянском языке, созданную для обучения моделей распознавания армянской речи и показать необходимость создания моделей распознавания речи как стратегически важные инвестиции.

Tigran Karamyan, Applications for automatic speech recognition in the economy and the necessity of creating an Armenian speech database – The use of speech recognition models is becoming more and more widespread, which not only makes life easier for people, but also makes it safer. Natural language processing techniques play a key role when working with unstructured data such as audio/text data. It is worth noting that natural language processing techniques can reveal whether sentiments, information, tone, or topics are associated with certain outcomes. These clues can be used to build testable hypotheses about what factors influence which decisions and how. On the other hand, these methods and their application in various industries make it possible to simplify the collection, processing, transmission and use of information.

The paper presents the main problems and difficulties in creating automated speech recognition models, especially for less common languages, as well as the application of speech recognition models in various sectors of the economy, such as finance, the service sector, healthcare, education, cyber security.

The goal is to present to the public the “Stories-15” database, which is a 15-hour audio / text data collected from about 40 stories in Armenian, created to train Armenian speech recognition models and show the need for creating speech recognition models as strategically important investments.

Ուղարկվել է խմբագրություն 02.05.2022թ.

Գրախոսվել է 16.05.2022թ.

Ստորագրվել է տպագրության 17.06.2022 թ.