



ПРЕПОЗНАВАЊЕ ЉУДСКИХ АКТИВНОСТИ УПОТРЕБОМ МОДЕЛА МАШИНСКОГ УЧЕЊА

HUMAN ACTIVITY RECOGNITION USING MACHINE LEARNING ALGORITHMS

Теодора Недић, Јелена Сливка, Факултет техничких наука, Нови Сад

Област – РАЧУНАРСТВО И АУТОМАТИКА

Кратак садржај – Проблем препознавања људских активности се користи за дефинисање образаца понашања човека. За препознавање се користе сигнали уграђених сензора паметних телефона или носивих уређаја. Сирове податке потребно је претпроцесирати применом филтера за шум и затим узорковати помоћу фиксираних клизајућих прозора. Из сваког прозора, вектор обележја се добија рачунањем статистичких променљивих из домена времена и фреквенције. У овом раду су, над тако обрађеним подацима, обучени статистички модели машинског учења и конволуционе неуронске мреже. Најбоље резултате постигли су модели насумичне шуме (F -мера 0.93) и конволуциона неуронска мрежа (F -мера 0.94). На основу добијених резултата и радова других аутора, сматра се да ефикасност различитих класификатора у највећој мери зависи од самог скупа података, због чега се у литератури може наћи велики број различитих техника за решавање овог проблема које дају сличне резултате.

Кључне речи: препознавање људске активности, конволуционе неуронске мреже, метод насумичне шуме, машинско учење.

Abstract – The problem of recognizing human activities is used to define patterns of human behavior. Signals from built-in sensors of smartphones or wearable devices are used for recognition. The raw data must be preprocessed using a noise filter and sampled using fixed sliding windows. A feature vector is obtained from each window by computing statistical variables from the time and frequency domain. In this paper, traditional machine learning and convolution neural network models are trained on the preprocessed data. The best results were achieved using the Random Forest classifier (0.93 F -measure) and the convolutional neural network (0.94 F -measure). Based on the obtained results and the surveyed literature, the efficiency of different classifiers largely depends on the used data set, which is why many different techniques yield similar results.

Keywords: activity recognition, CNN, Random Forest, machine learning

НАПОМЕНА:

Овај рад проистекао је из мастер рада чији ментор је била др Јелена Сливка, ванр. проф.

1. УВОД

У овом раду описан је проблем препознавања активности човека (енг. *Human Activity Recognition*, HAR) на основу података прикупљених помоћу сензора паметних мобилних телефона. Доступност сензора у мобилним платформама је омогућила развој низа практичних апликација за неколико научних области као што су здравство, *IoT* и паметни градови, безбедност и превоз. Сензори носивих уређаја (акцелерометар, жироскоп, оријентациони сензори) прикупљају сирове сигнале временских серија.

У овом раду је описано решење за проблем препознавања активности човека. Циљ је детектовати шест активности: лежање, седење, стајање, ходање, ходање уз степенице и ходање низ степенице. Решење је засновано на спајању три јавно доступна скупа података која се типично користе за дати проблем. Над прикупљеним подацима је обучено више модела машинског учења у циљу одређивања најефикаснијег модела. Коначно, прикупљеним подацима је придодат скуп података креиран од стране аутора овог рада.

Прегледом постојећих решења, примећује се употреба и неуронских мрежа и традиционалних метода машинског учења. Најбољи резултати пронађени у литератури су следећи: конволуциона неуронска мрежа (*Convolutional Neural Network*, CNN) је постигла скоро савршене резултате за покретне активности (99.66%), највише на класама као што су ходање уз степенице и низ степенице. Међутим, најлошији резултат исте мреже је постигнут за активност лежања (87.71%). Са друге стране, употребом метода потпорног вектора (*Support Vector Machine*, SVM) постигнут је бољи резултат за стационарне активности (94.91%) [1].

Новина овог рада у односу на претходна решења представља скуп података, који је комбинација више најчешће коришћених скупова. Над таквим скупом, најбољи резултат су постигли насумичне шуме (0.93 F -мера), и конволуциона неуронска мрежа (0.94 F -мера). Остали модели су имали приближне перформансе (> 0.9 F -меру).

У поглављу 2 биће дат преглед постојећих решења. Детаљи решења и имплементације система су описани у поглављу 3, а поглавље 4 садржи анализу добијених резултата.

Последње, 5. поглавље, представља закључак рада.

2. ПРЕТХОДНА РЕШЕЊА

Претраживањем литературе везане за област HAR, дошли смо до неколико корисних радова описаних у наставку.

У раду [2] предложена је конволуциона неуронска мрежа за класификацију, након експеримената са пет модела: CNN, LSTM, BLSTM, MLP и SVM. Експерименти су извршени над два јавно доступна скупа података: UCI HAR [3] и PAMAP2 [4]. Циљна варијабла је активност, која је ограничена на 6 могућих вредности (*walking, walkingupstairs, walkingdownstairs, sitting, standing, laying*). Подаци су подељени на тренинг и тест скуп у односу 7:3. Додатно, подаци су прочитани коришћењем методе клизећег прозора са преклапањем 50%. Због другачије фреквенције узорковања за два скупа података, коришћена је *downsampling* метода да би подаци били конзистентни и лаки за поређење. Сви модели су тренирани са *learning rate*-ом 0.001 у 200 итерација. За метрике евалуације коришћени су: а метрике евалуације коришћени су: а метрике евалуације коришћени су: *accuracy, precision, recall, F1* мера и *confusion* матрице. CNN модел је постигао најбољи резултат са 93.2% тачности на UCI скупу података и 91% на PAMAP2 скупу.

У раду [5] коришћене су традиционалне методе машинског учења: стабло одлуке, метод насумичне шуме, ET, метод *K* најближих суседа, логистичка регресија, SVM. Предложена метода је логистичка регресија са постигнутом тачношћу од 98.4%. Скуп података је јавно доступни скуп UCI HAR као и у раду [2]. Из рада [5] закључено је да су постигнути бољи резултати у односу на CNN, па самим тим не смемо у старту одбацити статистичке моделе машинског учења.

У раду [1] описани су експерименти са дубоким неуронским мрежама. Показано је да је мала величина филтера сажимања ($1 \times 2 - 1 \times 3$) погоднија. CNN је такође постигао скоро савршену класификацију у покретним активностима. Закључено је да су CNN бољи од *state-of-the-art* техника истраживања података у HAR-у за скуп података прикупљен од 30 волонтера (UCI HAR), постижући тачност од 94,79% на тест скупу са сировим подацима сензора, и 95,75% са додатним информацијама временске брзе Фуријеове трансформације. Подаци су стандардизовани и подељени у односу 7:3. У раду је детаљно описан експериментални приступ. Величина инпут вектора је 128, величина филтера је 1×3 , величина *pooling* филтера је 1×3 , активациона функција је *ReLU*, *learning rate* је 0.01, *dropout* вероватноћа 0.8, *minibatch* величина је 128, максималан број епоха 5000. За подешавање хипер-параметара користили су *greedy-wise* тренирање, где су подешавали број слојева, број *feature* мапа, величине конволуционих филтера и *pooling* филтера. CNN је постигла скоро савршене резултате за покретне активности (99.66%), највише на класама као што су *walking upstairs* и *walking downstairs*. Најлошији резултат је постигнут за *laying* (87.71%). Са друге стране, употребом SVM је постигнут бољи резултат за стационарне активности (94.91%), али, као већина других класификатора, није добро разликовао међусобно сличне активности, као што су ход уз и низ

степенице. Из овог рада је преузет начин подешавања хипер-параметара ради постизања бољих резултата.

3. МЕТОД

У наредним поглављу су описани скупови података коришћени у раду, начин на који је извршена експлоративна анализа података и архитектура испробаних модела за предикцију.

3.1. Скупови података

У пројекту су спојени следећи скупови података:

UCI HAR [3] обухвата мерење шест различитих дневних активности које је извршило 30 учесника користећи *Samsung Galaxy S II* телефон. Скуп садржи 560 вредности које представљају различите карактеристике активности, укључујући временске и фреквенцијске променљиве. Садржи укупно 10 299 инстанци.

UCI HAR AAL [6] скуп података представља додатак претходном, са циљем побољшања тачности алгоритама за препознавање понашања. Подаци су прикупљени од 30 учесника који су носили уређај за мерење око струка, а њихова старосна група се кретала од 22 до 79 година.

PAMAP2 [4] скуп података обухвата 3 850 505 инстанци различитих физичких активности, укључујући и шест основних активности које се налазе и у првом скупу података. Подаци су прикупљени од стране 9 учесника који су носили три уређаја за мерење на различитим деловима тела. Садржи 54 карактеристике, од којих су најзначајније оне прикупљене помоћу акселерометра и жироскопа.

Ручно креиран скуп података је настао мерењем сигнала апликацијом *Physics Toolbox Suite*[7] док су учесници изводили шест различитих активности, свака у трајању од пет минута. Телефон је био постављен у пределу струка, а након мерења, апликација је генерисала CSV датотеке са прикупљеним подацима. На основу тих датотека, креиран је скуп података који је прошао кроз исти процес претпроцесирања као и претходни скуп података (PAMAP2).

3.2. Обрада података

Процес претпроцесирања података захтевао је низ корака којим се из сирових вредности сигнала који се састоји из три осе, добија 560 обележја у домену времена и фреквенције. У зависности од скупа података, процес је измењен како би се прилагодио подацима.

Прво су избачене непотребне колоне, тј. оне колоне које се не постоје у прва два скупа (временско обележје, притисак, вредности уређаја који су ношени око зглоба ноге и око руке). Затим су избачени и редови који не садрже неке од шест активности, као и редови са недостајућим вредностима јер их није било много. Након тога смањена је фреквенција са 100 на 50 Hz, избацивањем сваког другог реда. Затим су подаци филтрирани помоћу *median* филтера и *low-pass Butterworth* филтера трећег реда са угаоном фреквенцијом 20 Hz како би се уклонио шум.

Сигнал убрзања је раздвојен на два сигнала, убрзање тела и гравитационо убрзање употребом другог *low-*

pass Butterworth филтера са угаоном фреквенцијом 0.3 Hz. Након тога, сигнал убрзања тела и угаона брзина су коришћени за добијање сигнала трзаја (енг. *Jerk* сигнал).

Магнитуда ових сигнала израчуната је помоћу Еуклидске удаљености. На крају, FFT (енг. *Fast Fourier Transform*) је примењен на неке од добијених сигнала.

Над претходно описаним скуповима података извршена је експлоративна анализа и обрада. Овај процес је врло често круцијалан у поступку предикције, јер је то једини начин да сирови подаци постану употребљиви. Први корак представља анализа и обрада сваког од скупова података понаособ.

3.3. Архитектура решења

У наставку су описане архитектуре коришћених метода и мрежа које су примењиване над спојеним скупом података.

Прва испробана метода је метода потпорних вектора (енгл. SVM - *Support Vector Machine*) која показује добре резултате на високо-димензионалним скуповима података, што одговара проблему који решавамо. Испробане су две верзије SVM алгоритма, *Linear* и *Kernel*.

У случају употребе *Linear SVM* алгоритма, као класификатор коришћена је *LinearSVC* метода из *sklearn* библиотеке са следећим параметрима:

- C у опсегу од 1-12 са кораком два,
- *tolerancy* (критеријум заустављања) са вредношћу 0.00005.

Изабрана вредност параметра C показала је оптималне перформансе на нашем скупу података. Подешавање хипер-параметара одрађено је применом *RandomizedSearchCV* методе у шест итерација.

Визуелизацијом обележја у 2D простору, утврђено је да подаци нису линеарно сепарабилни, те је из тог разлога испробана и *kernel SVM* метода. Како немамо предзнање о подацима, коришћен је RBF кернел, са следећим параметрима:

- $\gamma = 0.0078125$
- $C = 8$

На нашем примеру, најбољи резултат дало је мало гама са малом толеранцијом на грешке.

Следећи експеримент је био са моделом насумичне шуме. Коришћена је *RandomForestClassifier* класа *sklearn* библиотеке са следећим параметрима:

- $n_estimators = 201$
- $max_depth = 30$,

где $n_estimators$ представља број стабала.

Конволуциона мрежа је препоручена за временски или просторно повезане податке, па је испробана у последњем експерименту. CNN 1D мрежа се састоји из 2 конволуциона и 2 *fully connected* слоја. Између првог и другог је употребљен један *max pooling* слој. Након сваког слоја је примењена *batch* нормализација и *dropout* са параметром 0.5. Активациона функција је *ReLU*, осим у последњем слоју где је коришћен *Softmax*. Архитектура мреже приказана је на слици 1.

Model: "sequential_5"

Layer (type)	Output Shape	Param #
conv1d_12 (Conv1D)	(None, 138, 96)	1152
batch_normalization_10 (Batch Normalization)	(None, 138, 96)	384
dropout_14 (Dropout)	(None, 138, 96)	0
max_pooling1d_5 (Max Pooling 1D)	(None, 68, 96)	0
conv1d_13 (Conv1D)	(None, 68, 32)	15392
batch_normalization_11 (Batch Normalization)	(None, 68, 32)	128
dropout_15 (Dropout)	(None, 68, 32)	0
flatten_4 (Flatten)	(None, 2176)	0
dense_8 (Dense)	(None, 4096)	8916992
dropout_16 (Dropout)	(None, 4096)	0
dense_9 (Dense)	(None, 6)	24582

 Total params: 8,958,630
 Trainable params: 8,958,374
 Non-trainable params: 256

Слика 1. Архитектура CNN

4. РЕЗУЛТАТИ И ДИСКУСИЈА

Над скуповима података предвиђана је људска активност, сврстана у једну од следећих категорија: лежање, седење, стајање, ходање, ходање уз степенице и ходање низ степенице.

Експеримент је постављен поделом на тренинг, валидациони и тест скуп у размери 70:20:10. Подаци су подељени насумичним стратификованим узорковањем. За утврђивање тачности модела коришћене метрике су: *F*-мера (*F-measure*), тачност (*accuracy*), прецизност (*precision*) и одзив (*recall*).

У табели 1 су приказани сумирани резултати експеримената над тест подацима. Као најбољи од модела машинског учења показао се *Random Forest* алгоритам. Разлог његове примене било је превазилажење недостатака *Decision Tree* алгоритма. Оптимизацијом *Random Forest* алгоритма, добијена је тачност 93.8%, што је чини овај модел скоро подједнако добрим као и CNN.

	макро F-мера	тачност
<i>Decision Tree</i>	0.8690	0.8692
<i>Linear SVM</i>	0.9126	0.9047
<i>Kernel SVM</i>	0.9264	0.9161
<i>Random Forest</i>	0.9368	0.9382
CNN	0.9407	0.9305

Табела 1. Резултати експеримената

Конволуциона неуронска мрежа постигла је најбоље резултате, што је очекиван резултат на основу литературе прегледане у поглављу 2. Резултати постигнути на тест скупу показују приближну тачност за све три коришћене метрике, од 94%. У табели 2 се налази преглед резултата по класама.

	прецизност	одзив	F-мера
walking	0.9490	0.9411	0.9450
walking_downstairs	0.9384	0.9366	0.9375
walking_upstairs	0.9432	0.9351	0.9391
laying	0.9465	0.9499	0.9482
sitting	0.9245	0.9164	0.9204
standing	0.9153	0.9072	0.9112

Табела 2. Резултати по класама за CNN модел

Можемо приметити да су модели најлошије препознавали активност стајања. Претпоставља се да је разлог томе сличност ове класе са класом седење.

Резултати по класама су у великој мери међусобно слични, док у постојећим решењима постоје веће варијације између класа. Претпоставка је да ово запажање може бити последица употребе више јавно доступних скупова података који су мерени различитим уређајима и у условима који нису потпуно исти. Такође, применом истог модела само над ручно прављеним скупом података, добијена је висока тачност од 99%, што даје могућност за будуће проширење скупа, укључивањем више учесника у процес мерења и повећањем броја активности.

Претпоставља се да би се ограничења могла превазићи и обучавањем комплекснијих модела неуронских мрежа, које захтевају много више ресурса у контексту хардвера.

5. ЗАКЉУЧАК

У раду је описан решавањем проблем препознавања људских активности на основу података измерених помоћу сензора акцелерометра и жироскопа.

Мотивацију за решавање овог проблема налазимо у могућности практичне употребе. Решење описано у овом раду заснива се на спајању три јавно доступна скупа података који се користе за дати проблем, где је за спајање било потребно претпроцесирати један од скупова како бисмо добили одговарајућа обележја. Затим следи обучавање неколико модела машинског учења како бисмо одредили оптималан модел и урадили поређење резултата. Након тога је ручно направљен скуп података, по узору на поступак за јавно доступне скупове. Најбољи добијени резултат је постигнут коришћењем CNN модела, где F-мера мерена на тест скупу износила 0.94. Одмах за њим је и модел насумичне шуме са F-мером 0.93. Додавањем ручно прављеног скупа података успешно су побољшани укупни резултати за 0.01 F-мере.

Квантитативно поређење HAR приступа омета чињеница да сваки систем ради са различитим скупом података [8]. Предност решења описаног у раду у односу на постојећа је процес спајања више извора података, који су мерени различитим уређајима и у условима који нису потпуно идентични. Претпоставља се да су због тога резултати нешто нижи у односу на најбоље резултате пронађене у литератури, али је модел истрениран тако да је робуснији и може се користити за препознавање активности из сва три коришћена скупа података. За сада је решаван проблем препознавања шест активности, али, узимајући у обзир добре постигнуте резултате, може се у даљим истраживањима додати и низ других људских активности.

6. ЛИТЕРАТУРА

- [1] Rona, C. A., & Cho, S. B. (2016). Human activity recognition with smartphone sensors using deep learning neural networks. *Expert systems with applications*, 59, 235-244.
- [2] Wan, S., Qi, L., Xu, X., Tong, C., & Gu, Z. (2020). Deep learning models for real-time human activity recognition with smartphones. *Mobile Networks and Applications*, 25(2), 743-755.
- [3] UCI HAR dataset [Reyes-Ortiz, Jorge, Anguita, Davide, Ghio, Alessandro, Oneto, Luca & Parra, Xavier. (2012). Human Activity Recognition Using Smartphones. UCI Machine Learning Repository.]
- [4] PAMAP2 dataset [A. Reiss and D. Stricker. Introducing a New Benchmarked Dataset for Activity Monitoring. The 16th IEEE International Symposium on Wearable Computers (ISWC), 2012.]
- [5] Minarno, A. E., Kusuma, W. A., Wibowo, H., Akbi, D. R., & Jawas, N. (2020, June). Single triaxial accelerometer-gyroscope classification for human activity recognition. In 2020 8th International Conference on Information and Communication Technology (ICoICT) (pp. 1-5). IEEE.
- [6] Smartphone Dataset for Human Activity Recognition (HAR) in Ambient Assisted Living (AAL) Data Set
- [7] Vieyra, Chrystian, Vieyra, Rebecca. Physics Toolbox Suit
- [8] Lara, O. D., & Labrador, M. A. (2013). A Survey on Human Activity Recognition using Wearable Sensors. *IEEE Communications Surveys & Tutorials*, 15(3), 1192-1209. doi:10.1109/SURV.2012.110112.00192

Кратка биографија:



Теодора Недић рођена је 1998. године у Врбасу, Србија. Основне академске студије завршила је 2021. године на Факултету техничких наука, на ком брани и мастер рад 2023. године из области Електротехнике и рачунарства – Софтверско инжењерство и информационе технологије. контакт: nedicteodora1@gmail.com