AUTHOR(S): SANI, S., WIRATUNGA, N., MASSIE, S. and COOPER, K.

TITLE: Study of similarity metrics for matching network-based personalised human activity recognition.

YEAR: 2018

Publisher citation: SANI, S., WIRATUNGA, N., MASSIE, S. and COOPER, K. 2018. Study of similarity metrics for matching network-based personalised human activity recognition. In Minor, M. (ed.) Workshop proceedings of the 26th International conference on case-based reasoning (ICCBR 2018), 9-12 July 2018, Stockholm, Sweden. Stockholm: ICCBR [online], pages 91-95. Available from: http://iccbr18.com/wp-content/uploads/ICCBR-2018-V3.pdf#page=91

OpenAIR citation: SANI, S., WIRATUNGA, N., MASSIE, S. and COOPER, K. 2018. Study of similarity metrics for matching network-based personalised human activity recognition. In Minor, M. (ed.) Workshop proceedings of the 26th International conference on case-based reasoning (ICCBR 2018), 9-12 July 2018, Stockholm, Sweden. Stockholm: ICCBR, pages 91-95. Held on OpenAIR [online]. Available from: https://openair.rgu.ac.uk/

Study of Similarity Metrics for Matching Network-Based Personalised Human Activity Recognition

Sadiq Sani[1], Nirmalie Wiratunga[1], Stewart Massie[1], and Kay Cooper[2]

[1] School of Computing Science and Digital Media,
[2] School of Health Sciences,
Robert Gordon University,
Aberdeen AB10 7GJ, Scotland, UK
{s.sani,n.wiratunga,s.massie,k.cooper}@rgu.ac.uk

Abstract. Personalised Human Activity Recognition (HAR) models trained using data from the target user (subject-dependent) have been shown to be superior to non personalised models that are trained on data from a general population (subject-independent). However, from a practical perspective, collecting sufficient training data from end users to create subject-dependent models is not feasible. We have previously introduced an approach based on Matching networks which has proved effective for training personalised HAR models while requiring very little data from the end user. Matching networks perform nearest-neighbour classification by reusing the class label of the most similar instances in a provided support set, which makes them very relevant to case-based reasoning. A key advantage of matching networks is that they use metric learning to produce feature embeddings or representations that maximise classification accuracy, given a chosen similarity metric. However, to the best of our knowledge, no study has been provided into the performance of different similarity metrics for matching networks. In this paper, we present a study of five different similarity metrics: Euclidean, Manhattan, Dot Product, Cosine and Jaccard, for personalised HAR. Our evaluation shows that substantial differences in performance are achieved using different metrics, with Cosine and Jaccard producing the best performance.

1 Introduction

Automatic recognition and tracking of human activity using wearable sensors is increasingly being adopted for health care applications e.g. management of chronic low back pain in SelfBACK [1] [1]. An important consideration for HAR applications is classifier training, where training examples can either be acquired from a general population (subject-independent), or from the target user of the system (subject-dependent). Previous works have shown using subject-dependent data to result in superior performance [2–4]. Matching networks [6]

[1]The SelfBACK project is funded by European Union's H2020 research and innovation programme under grant agreement No. 689043.

have been successfully applied for efficiently learning personalised HAR models [5]. Given a (typically small) support set of labelled examples, matching networks are able to classify an unlabelled example by reusing the class labels of the most similar examples in the support set. A key advantage of matching networks is that they use metric learning to produce feature embeddings or representations that maximise classification accuracy, given a chosen similarity metric. Thus, it is important to investigate how different similarity metrics affects the performance of matching networks for personalised HAR. Accordingly, in this paper, we present a study of five different similarity metrics used with matching networks.

2 Personalised HAR using Matching Networks

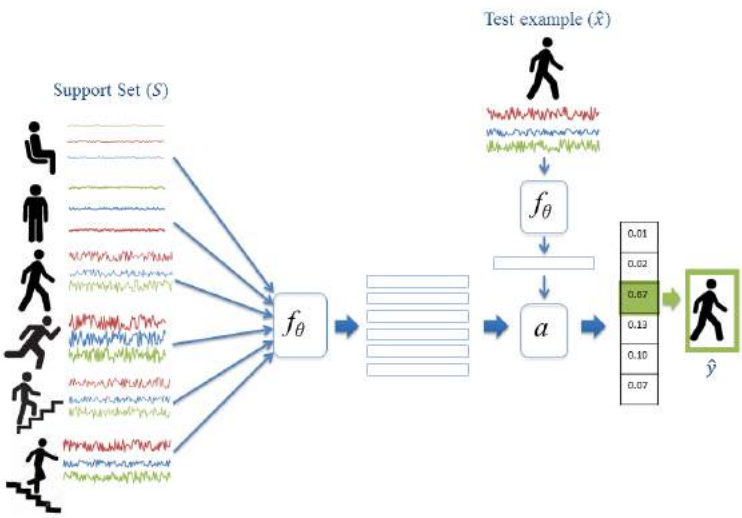

Fig. 1. Illustration of matching network for HAR.

The aim of matching networks is to learn a model that maps an unlabelled example $\hat{x}$ to a class label $\hat{y}$ using a small support set S of labelled examples. This is illustrated in Figure 1. Given a set of instances $X = \{x|x \text{ is an instance vector}\}$, a set of class labels $L = \{y|y \text{ is a class label}\}$, an embedding function f_θ which in this is case a neural network parameterised by ρ, the function a is an attention mechanism that takes the embedded representation of a test instance and a support set S and returns a probability distribution $P(y|\hat{x}, S)$ over class labels y of instances in S. To train the matching network for personalised HAR, we also define a set of users U where each user $u_j \in U$ is comprised of a set of labelled examples as follows:

$$u_j = \{(x, y)|x \in X, y \in L\} \tag{1}$$

Next we define a set of training instances T_j for each user u_j as follows:

$$T_j = \{(S_j, B_j)\} \tag{2}$$

i.e., T_j is made up of user-specific support and target set pairs S_j and B_j respectively, where $S_j = \{(x, y)|(x, y) \in u_j\}$ and $B_j = \{(x, y)|(x, y) \in u_j, (x, y) \in S_j\}$. Note that the set of labels in S_j is always equivalent to L because we are interested in learning a classifier over the entire set of activity labels. Accordingly, S_j contains m examples for each class $y \in L$ and the cardinality of S_j is $|S| = m \times |L|$. Both S_j and B_j are sampled at random from u_j l times to create T_j. Each B_j is used with it's respective S_j by classifying each instance in B_j using S_j and computing loss using categorical cross entropy. The network is trained using stochastic gradient descent and back propagation.

3 Similarity Metrics

Matching networks use a similarity metric to match a given test instance to the most similar instances in a support set. In the following subsections, We discuss five of the most popular similarity metrics used in literature.

3.1 Euclidean

Euclidean distance is perhaps the most popular metric used for estimating similarity between items represented as numerical vectors. The Euclidean metric gives the distance between any two points in n-dimensional space as the length of a straight line connecting those two points. Euclidean distance can be converted to similarity simply by taking the inverse as shown in Equation 3.

$$Euclidean(\hat{x}, x) = \frac{1}{\sum\sum (\hat{x}_j \sim x_j)^2 + 1} \tag{3}$$

3.2 Manhattan

The Manhattan distance between two items is computed as the sum of absolute differences between the values of their dimensions. This can also be converted to a similarity by taking the inverse as shown in Equation 4. In comparison with Euclidean, the Manhattan metric is less susceptible to large differences in values in few dimensions.

$$Manhattan(\hat{x}, x) = \frac{1}{\sum |\hat{x}_j \sim x_j| + 1} \tag{4}$$

3.3 Cosine

Cosine metric estimates similarity between two items by measuring the angle between their vectors in n-dimensional space. Cosine similarity can be computed as shown in Equation 5.

$$Cosine(\hat{x}, x) = \frac{\sum_j^n \hat{x}_j x_j}{\sqrt{\sum_j^n \hat{x}_j^2}\sqrt{\sum_j^n x_j^2}} \tag{5}$$

3.4 Dot Product

Dot product measures the projection of one vector onto another in n-dimensional coordinate space as shown in Equation 6. Unlike cosine similarity, dot product is not normalised and thus takes into account both angle and magnitude of the two vectors.

$$DotProduct(\hat{x}, x) = \sum \hat{x}_j x_j \tag{6}$$

3.5 Jaccard

The Jaccard metric measure similarity between finite sets as the ratio of the size of the intersection to the size of the union of the sets. The general form of the Jaccard metric for finding similarity between numerical vectors is provided in Equation 7

$$Jaccard(\hat{x}, x) = \frac{\sum \hat{x}_j x_j}{\sum \hat{x}_j^2 + \sum x_j^2 \sim \sum \hat{x}_j x_j} \tag{7}$$

4 Evaluation

Evaluation is conducted on a dataset of 50 users with 9 activity classes and about 3 minutes of activity data per class. We adopt a hold-out validation strategy where 8 out of the 50 users are randomly selected for testing. To simulate user provided samples for creating personalised support sets, we hold out the first 30 seconds of each test user's data for creating the support set. This leaves approximately 150 seconds of data per activity which are used for testing, Performance is reported using macro-averaged F1 score.

In the evaluation, we explore the performance of matching network for personalised HAR using the five similarity metrics presented in Section 3. Five different matching networks are trained, each using one of the similarity metrics. All matching networks are identical except for the similarity difference in similarity metric are all trained for the same number of epochs. Results are presented in Table 1. It can be observed that the best results are achieved using Cosine and Jaccard. Euclidean produces a reasonably close third place performance while both Manhattan and Dot Product are a distant forth and fifth place respectively. Note that both Cosine and Jaccard metrics are normalised by the magnitude of the vectors involved in the similarity. This suggests that similarity metrics that do not take into account differences in vector magnitudes tend to work better for this application.

Table 1. Results of different algorithms showing F1 scores.

Metric	Euclidean	Manhattan	Cosine	Dot Product	Jaccard
F1 Score	0.757	0.696	**0.788**	0.694	0.783

5 Conclusion

In this paper, we have presented a comparative study of 5 different similarity metrics for personalised HAR using matching networks. Results show cosine and Jaccard metrics to produce the best classification performance. Our work suggests that the choice of similarity metric is a very important consideration for matching networks and that the performance of metrics that do not consider differences in vector magnitude (e.g. Cosine and Jaccard) are superior to metrics that take into account vector magnitude (e.g. Euclidean, Manhattan and dot product).

References

1. Bach, K., Szczepanski, T., Aamodt, A., Gundersen, O.E., Mork, P.J.: Case representation and similarity assessment in the selfback decision support system. In: Proceedings of 24th International Conference on Case-Based Reasoning. pp. 32–46. Springer International Publishing (2016)
2. Berchtold, M., Budde, M., Gordon, D., Schmidtke, H.R., Beigl, M.: Actiserv: Activity recognition service for mobile phones. In: Proceedings of International Symposium on Wearable Computers. pp. 1–8 (2010)
3. Jatoba, L.C., Grossmann, U., Kunze, C., Ottenbacher, J., Stork, W.: Context-aware mobile health monitoring: Evaluation of different pattern recognition methods for classification of physical activity. In: Proceedings of 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. pp. 5250–5253 (2008)
4. Sani, S., Wiratunga, N., Massie, S., Cooper, K.: knn sampling for personalised human activity recognition. In: Proceedings of International Conference on Case-Based Reasoning. pp. 330–344. Springer (2017)
5. Sani, S., Wiratunga, N., Massie, S., Cooper, K.: Personalised human activity recognition using matching networks. In: Proceedings of International Conference on Case-Based Reasoning. Springer (2018)
6. Vinyals, O., Blundell, C., Lillicrap, T., Wierstra, D.: Matching networks for one shot learning. In: Proceedigns of Advances in Neural Information Processing Systems. pp. 3630–3638 (2016)