

Smartphone-Based Recognition of Human Activities and Postural Transitions

Bruno M. Dobrovolski¹, Eduardo A. Schmoller¹

¹Departamento Acadêmico de Informática – Universidade Tecnológica Federal do Paraná
Pato Branco – PR – Brasil

{brunod,schmoller}@alunos.utfpr.edu.br

1. Introdução

Exemplo artigo: <https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2013-11.pdf>

Este trabalho apresentará resultados da aplicação de métodos de aprendizado de máquinas na predição de movimentos humanos. Foram implementados os seguintes métodos de aprendizado: KNN, Perceptron e SVM.

Os dados para base para treinamento e teste foram obtidos de [Dheeru and Karra Taniskidou 2017]. Essas informações são referente a sensores de um dispositivo móvel Galaxy Samsung S2. O *smartphone* foi mantido junto ao corpo dos voluntários em atividades de: Levantar, sentar, deitar, caminhar, subir escadas e descer escadas. Os testes foram executados com 30 voluntários e os dados coletados através dos sensores do smartphone, acelerômetro de 3 eixos e giroscópio de 3 eixos.

Os dados obtidos dos sensores foram pré-processados com a aplicação de filtros para remoção de ruídos, amostrados em intervalos constantes de tempo e normalizados. Também foram segmentados separando cada um dos 3 eixos e calculando várias informações estatísticas (totalmente descritas em citeDua:2017), como média, desvio padrão, máximo, mínimo, entre outras, totalizando 561 características. Cada conjunto de dados que representa os movimentos recebeu uma identificação. O conjunto de características e de identificações (*label*) são entradas para os algoritmos de classificação.

Como forma de validação do processo de aprendizado, o conjunto de dados de entrada foi dividido em dois subconjuntos denominados conjunto de treino e conjunto de testes. O conjunto de treino é utilizado no processo de treinamento, após o processo de treinamento o conjunto de testes é classificado e os resultados são comparados com os resultados esperados.

Ainda, os algoritmos foram alterados através da redução de dimensionalidade com o uso do método PCA. Como esse método não entrega um resultado ótimo, foi necessário variar o número de componentes para encontrar a melhor combinação.

Como métrica para a determinação da qualidade do processo de aprendizados foram utilizadas a acurácia da predição e utilização do *kappa score*. O *kappa score* funciona como uma validação do resultado apresentado pela acurácia, visto que ele considera em seu calculo os resultados falso-positivos e falso-negativos, colocando peso em cada uma das características.

2. Revisão

Esta seção apresenta os métodos aplicados para a classificação dos movimentos.

2.1. Perceptron

http://ml.informatik.uni-freiburg.de/former/_media/documents/teaching/ss09/ml/perceptrons.pdf

Classificador linear

2.2. KNN

KNN algoritmo simples que realiza o aprendizado de todas as entradas e classifica as novas entradas com base na similaridade dos casos já aprendidos. Ao realizar a classificação de uma nova entrada é retornado k vizinhos mais próximos, classificação se dá pela maioria dos vizinhos.

2.3. SVM

SVM é um classificador definindo a separação entre as classes através de um hiperplano, a entrada são dados com *labels* e a saída é um hiperplano ótimo para categorização de novas entradas. Em duas dimensões o hiperplano é representado através de uma reta que divide o plano em duas partes.

2.4. PCA

Redutor de dimensionalidade selecionando as principais componentes, entre as formas de encontrar as componentes principais é buscar as projeções que maximizem a variância.

2.5. Testes Estatísticos

Testes estatísticos são utilizados para validação de hipóteses usando dados amostrais para decidir sobre a validade da hipótese.

Cenários típicos de uso de testes estatísticos em métodos de aprendizado: adição de etapas de processamento, comparar diversos algoritmos, validação de capacidade de generalização de um classificador. O princípio básico é a formulação de hipóteses e decidir se os dados mostram alguma evidencia para rejeição da hipótese.

2.5.1. Validação Cruzada

<https://www.analyticsvidhya.com/blog/2018/05/improve-model-performance-cross-validation-in-python-r/>
<https://towardsdatascience.com/cross-validation-in-machine-learning-729>
<https://www.kaggle.com/dansbecker/cross-validation>
https://www.researchgate.net/profile/Ron_Kohavi/publication/2352264_A_Study_of_Cross-Validation_and_Bootstrap_for_Accuracy_Estimation_and_Model_Selection/links/02e7e51bcc14c5e91c000000.pdf
http://scikit-learn.org/stable/modules/model_evaluation.html

2.5.2. Matriz de Confusão

http://www2.cs.uregina.ca/~hamilton/courses/831/notes/confusion_matrix/confusion_matrix.html

2.5.3. Kappa Score

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900052/>
<http://www.pmean.com/definitions/kappa.htm>
<http://www.statisticshowto.com/cohens-kappa-statistic/>
<https://academic.oup.com/ptj/article/85/3/257/2805022>

3. Resultados

Com a aplicação dos métodos citados anteriormente foi possível apresentar a Figura 1. Esse gráfico representa a busca pelo melhor agrupamento de características utilizando PCA. Como é possível observar, cada um dos métodos tem um comportamento diferente. O Perceptron tem melhores resultados quanto maior for a quantidade de características; Já o KNN não apresenta muitas melhoras; Por fim o SVM tem uma redução da acurácia para maior quantidade de características. É importante notar que apesar da variação nos resultados, as perdas são relativamente pequenas, cerca de 5%. O código executado partiu de 10 grupos a 260 grupos, com incremento de 10. Porém, é sabido que os bons resultados

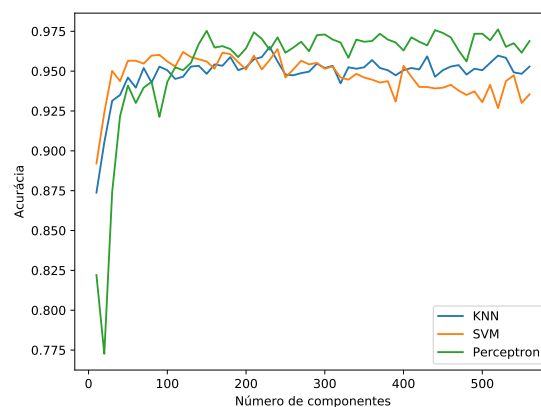
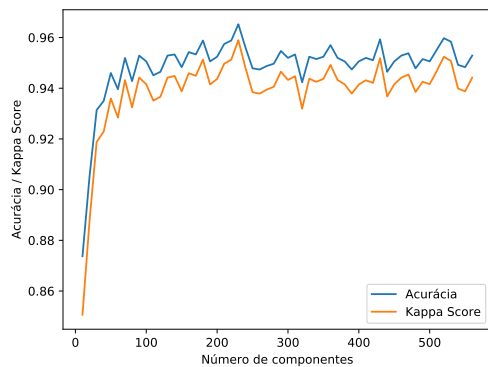
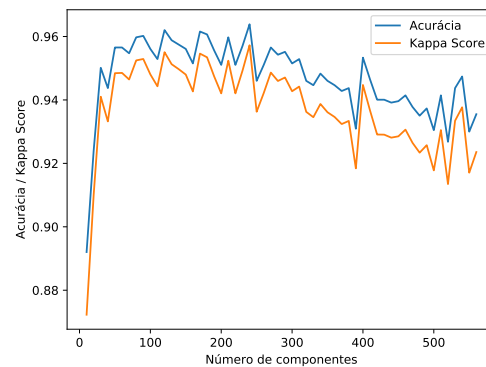


Figura 1. Acurácia - KNN, SVM, Perceptron

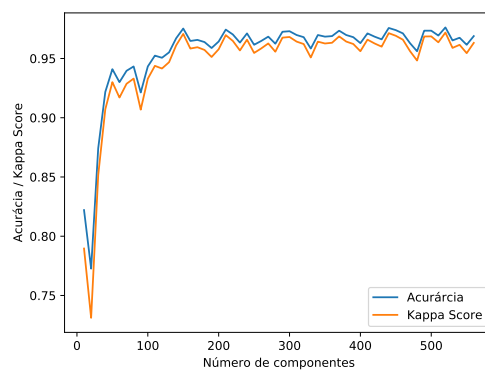
da acurácia podem não ser reais. Para isso, foi aplicado o Kappa-Score, como explicado anteriormente ele pondera os falso-positivos e falso-negativos, quando comparado à acurácia a diferença entre esses dois valores informa a qualidade do método. A Figura 2, apresenta essa comparação em relação aos 3 métodos.



(a) Comparação - KNN



(b) Comparação - SVM



(c) Comparação - Perceptron

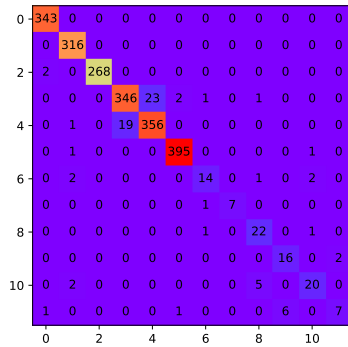
Figura 2. Comparação entre Acurácia e Kappa Score

Observando esses resultados, foi possível obter o melhor agrupamento para cada um dos métodos, essa escolha é apresentada na Tabela 1.

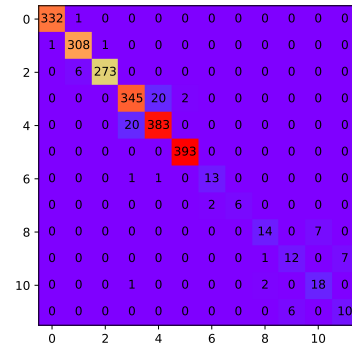
Tabela 1. Resultados

Método	Acurácia	Kappa Score	Agrupamento
KNN	0.9652	0.9589	230
Perceptron	0.9762	0.9718	520
SVM	0.9663	0.9572	240

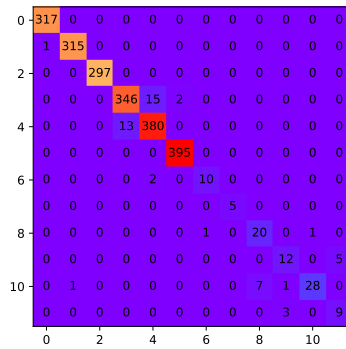
Para o caso ótimo desses conjuntos foi construída a matriz de confusão, afim de validar as informações apresentadas pela acurácia e Kappa Score. As matrizes apresentadas na Figura 3



(a) Matriz de Confusão - KNN



(b) Matriz de Confusão - SVM



(c) Matriz de Confusão - Perceptron

Figura 3. Comparação entre Matrizes de confusão

4. Conclusões

Este trabalho permitiu a compreensão de alguns métodos de aprendizagem. A escolha de cada método deve ser adequada ao tipo dos dados a serem aprendidos e, neste caso, a não uniformidade da quantidade de informações sobre cada característica gerou uma diferença considerável entre cada método. Também notou-se que a técnica PCA não necessariamente colabora com os resultados, por exemplo, no método Perceptron, quanto mais características, melhor foram os resultados. Outro fator notável, foi o ganho ao aumentar a quantidade de classes, isso gerou um ganho relativamente pequeno de acurácia e aumentou o tempo de computação consideravelmente. De forma conclusiva, não houve um método com destaque nos resultados obtidos, as diferenças foram sutis e todos foram considerados satisfatórios.

Referências

[Dheeru and Karra Taniskidou 2017] Dheeru, D. and Karra Taniskidou, E. (2017). {UCI} machine learning repository.