Human Activity Recognition using Smartphone Sensors

Jessica Moore, Binghai Ling
jmoore5@stanford.edu, binghai@stanford.edu

Summary

- Cell phones are becoming ubiquitous and can reasonably be used to track daily activity
- We predicted the activity a person was taking, based on signals from their cell phone
- Using PCA and SVMs, we achieved a ~4% error rate with only about 200 features
- Most errors in the final models come from confusion of the ‘sitting’ and ‘standing’ classes

Data

- Source: UCI Machine Learning Repository
- Signals from waist-worn cellphones of 30 individuals performing 6 activities
- Features are summary statistics (e.g., mean, SD) of filtered accelerometer and gyroscope signals, obtained during a period of activity

Feature Engineering

Due to the size of the data set, variance seems likely to be the primary source of prediction error. To combat this, we implemented two unsupervised feature engineering methods:

- **Principal Components Analysis**

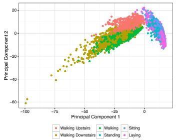
Eigenvalue decomposition of:

$$\Sigma = \frac{1}{m}\sum_{i=1}^{m} x^{(i)}x^{(i)^T}$$

Features are scores:

$$Z_{ij} = u_j^T x^{(i)}$$

where u_j is the j-th PC

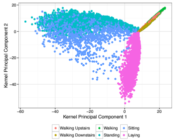


- **Kernelized Principal Components Analysis**

Implicit eigenvalue decomposition of:

$$\Sigma = \frac{1}{m}\sum_{i=1}^{m} \phi(x^{(i)})\phi(x^{(i)})^T$$

Accomplished via an eigenvalue decomposition of the kernel matrix



Features are scores:
Projections of data onto kPCs

Models

- **K-Nearest Neighbors**

Classification based on k nearest training points

$$P(y^{(i)} = j \mid x = x^{(i)}) = \frac{1}{k}\sum_{q\in N} 1\{y^{(q)} = j\}$$

- **Softmax Regression**

Classification based on maximum $\theta_j^T x^{(i)}$

$$P(y^{(i)} = j \mid x = x^{(i)}) = \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^{L} e^{\theta_l^T x^{(i)}}}$$

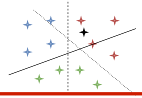
- **Support Vector Machines**

Find α to minimize

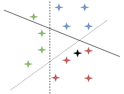
$$J(\alpha)_\lambda = \frac{1}{m}\sum_{i=1}^{m}\left[1 - y^{(i)}K^{(i)}\alpha\right] + \frac{\lambda}{2}\alpha^T K\alpha$$

where K is a kernel matrix. We use a linear kernel: $x^T y$

One vs. one



One vs. all



Results

Test Error

	PC Features	kPC Features	Original Data
kNN	0.127 (290 PCs)	0.171 (380 kPCs)	0.098
Softmax	0.052 (299 PCs)	0.129 (212 kPCs)	0.109
SVM (ovo)	0.039 (220 PCs)	0.231 (390 kPCs)	0.039
SVM (ova)	0.039 (204 PCs)	0.165 (397 kPCs)	0.038

Training: 5,147 obs.; Validation: 2,206 obs.; Test: 2,947 obs.

Features:

- kPC-based features yielded high errors
- PC-based features gave similar errors to the full, original data
- SVMs performed well and used only ~200 PCs

Algorithms:

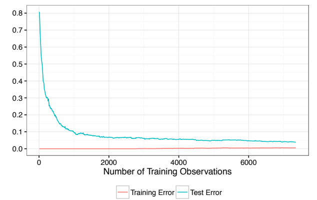
- kNN performed significantly worse than Softmax regression and the two SVM methods
- Both SVM methods out-performed Softmax regression, obtaining a ~4% error rate

Error Analysis & Future Work

Predicted

		WU	WD	Walk	Stand	Sit	Lay
Actual	**WU**	454	0	17	0	0	0
	WD	5	410	3	0	2	0
	Walk	4	0	491	1	0	0
	Stand	0	0	0	514	18	0
	Sit	1	0	1	64	425	0
	Lay	0	0	0	0	0	537

- Errors primarily from misclassification of ‘sitting’ and ‘standing’ observations
- Future work might include additional methods to differentiate those classes



- Low training error and distinctly higher test error suggests high variance
- Future work might include regularization or additional feature reduction to resolve this

Note: Results above are for one vs. one SVMs; the results for one vs. all SVMs are nearly identical.

Discussion

- Using PC-based features dramatically reduced the number of features, without affecting misclassification error.
- The success of SVMs and Softmax regression (with PC-based features) is likely because the classes are separable in the space of the first ~200-300 principal components.
- kPCA failure may indicate a lack of relationship between nonlinear variation and the response. This is surprising, given that prior authors had success with Gaussian-kernel SVMs.
- Primary errors in the final model arise from mislabeling of ‘sitting’ and ‘standing’ classes. These are often confused for one another.

Acknowledgements

We would like to thank Professor Ng and Professor Duchi for their instruction throughout the quarter. We would also like to thank our Project TA, Francois, for his comments as we completed this project.

References:

1) B. Rose, No, Phones Aren't More Accurate Than Fitness Wearables. Wired. (Mar. 18, 2015).
2) D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, eds., A Public Domain Dataset for Human Activity Recognition Using Smartphones., 21th Europeo Symposium on Articial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013. (Bruges, Belgium, 2013).
3) L. Bao and S. S. Intille, Activity Recognition from User-Annotated Acceleration Data. PERVASIVE 2004, LNCS 3001.
4) O. D. Lara and M. A. Labrador, IEEE Communications Surveys & Tutorials: A Survey on Human Activity Recognition using Wearable Sensors (Third Quarter 2013).
5) D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, Human Activity Recognition on Smartphones Using a Multiclass Hardware-Friendly Support Vector Machine. In Ambient Assisted Living and Home Care (2012).
6) D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, ESANN 2013 proceedings, European Symposium on Articial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) (24-26 April 2013).
7) T. H. G. James, D. Witten and R. Tibshirani, Introduction to Statistical Learning with Applications in R, 6th ed. (Springer, 2013).
8) A. Ng and J. Duchi, CS229: Machine Learning - Course Notes. (Autumn 2016).
9) K.-R. M. B. Scholkopf, A. Smola, Kernel Principal Components Analysis.
10) S. Raschka, Kernel Tricks and Nonlinear Dimensionality Reduction via RBF Kernel PCA (Sept. 14, 2014).
11) M. Welling, Kernel Principal Components Analysis.
12) R. Osadchy, Unsupervised Learning 2011: Kernel PCA (2011).
13) L. Mackey, STATS202: Data Mining and Analysis - Course Notes. (Autumn 2015).
14) K. Ganesan, Computing Precision and Recall for Multi-Class Classication Problems.
15) E. P. L. van der Maaten and J. van den Herik, Dimensionality Reduction: A Comparative Review (October 26, 2009).
16) T. Hastie, R. Tibshirani, and J. Friedman, Elements of Statistical Learning, 3rd ed. (Springer, 2013).