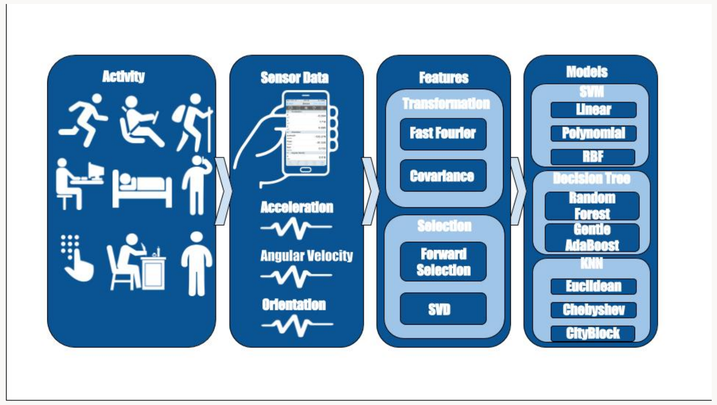
Human Activity Recognition via Cellphone Sensor Data

Wei Ji, Heguang Liu, Jonathan Fisher
wxj@stanford.edu, heguangl@stanford.edu, jofisher@stanford.edu
Stanford University

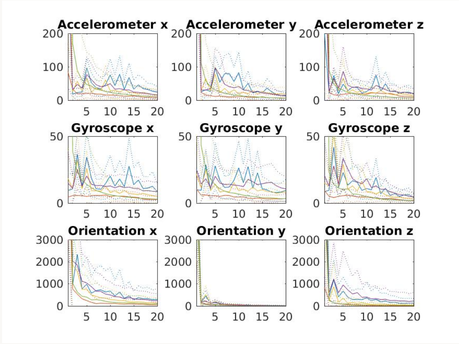
Overview

The purpose of this project is to identify human activities while using cell phones via mobile sensor data. We collect 2085 data samples, which includes 3-axis acceleration, angular velocity and orientation sensor data, from 4 volunteers using MATLAB Mobile package. After cleaning, interpolating, FFT, we get 135 raw features, and we further reduce the feature number to 21 via feature selection. After comparing the results of different models, eg Decision Tree, SVM, KNN, we successfully build an Ensembled Bagged Trees Model which gives 95.7% prediction accuracy over 626 test data on 9 human activities(walking, running, driving, typing, writing, talking, laying, sitting, standing).



Data

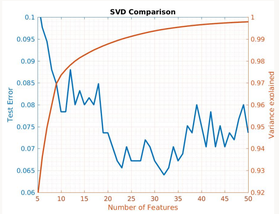
- Data Collection
  We collected 60M, approximately 15000 seconds, data at 50Hz sample rate from 4 volunteers using MATLAB Mobile, Android/iOS sensors support package and Mathworks Cloud.
  - Connect Use MATLAB Mobile App on the mobile device to connect to MathWorks Cloud, load collecting script, and tag activities.
  - Record Run the scripts which initialize a mobiledev object, enable sensors, set sample rate to 50Hz and start recording timer.
- Data Processing
  - Segment We divide each recording sample into 128*0.02=2.56 second segments. And each segment is used as a data point. We also remove duplicate, missing data.
  - Interpolate Since sensors are not exactly synchronized, and time internal between collected samples is in a range of [0.018, 0.022] second, we use Linear Grid Interpolation to get synchronized data point.



Features

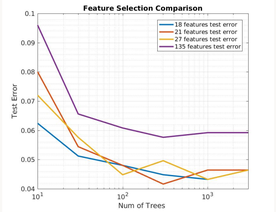
- Raw Features
  - 9*128 raw features from 9 sensors on time domain.
- Features Transformation
  - Covariance matrix of every 2 sensors gives 45 features.
  - First 10 spectrum in frequency domain show high coefficients, gives 9*10 features.
- Feature Selection
  - 135 features is still too large to make algorithms run efficiently. We adopt SVD and forward feature selection to reduce feature dimensions.
  - SVD reduce feature dimension

    Running a serious of SVD on a bagged tree model shows that Top 7 features can preserve 95% of total energy, and give a 9.3% test error. Top 22 features can preserve 99% of total energy, which gives 6.6% test error.



  - Forward selection reduce feature dimension

    Running series of forward feature selection on the bagged tree model shows that, 18, 21, 27-feature model overplay 135 feature model. The 21 feature model achieves lowest test error 4.3% at number of trees larger than 300.

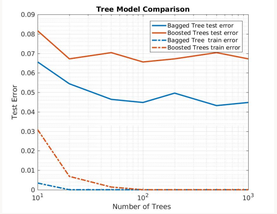


Models Comparison

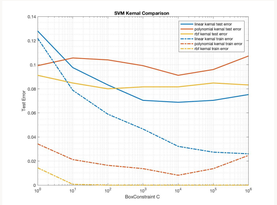
This is a standard Supervised Learning problem to find the best model to predict 9 labels with 21 features. We tried to solve this problem using Decision Tree, SVM and KNN.

- Decision Tree Comparison

  We use bagged tree (Random Forest), Boosted Tree (Gentle Adaboost) with different number of trees. Results show that bagged trees with 300 trees achieves the lowest error rate 4.3%.

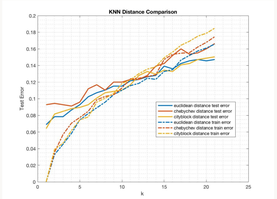


- SVM Comparison
  We tried Linear, Polynomial and RBF(Gaussian) Kernel, using L1-regulation with various box constrains. Results show that linear kernel achieves lowest error rate at 6.8% with C=10000. RBF, Polynomial kernal models have the worse results.



- KNN Comparison

  We tried Euclidean, Chebychev and Cityblock distance function to measure the KNN distance with various k. Results show that KNN has a worse performance comparing with Tree and SVM Models.

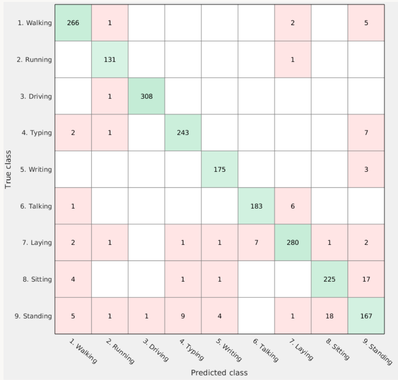


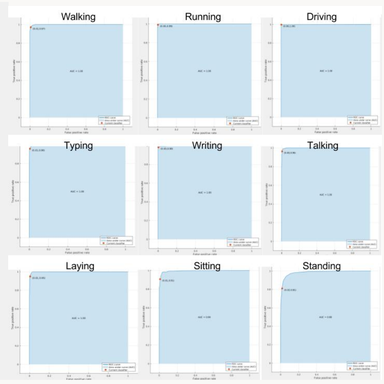
Results and Discussion

Based on the model comparison, we find the optimized algorithm is Bagged Trees with Number of trees bigger than 300. The Bagged Tree model is built using formula (1). Based on the Confusion Matrix and ROC, we discovered the following facts:

- Less intensive activities, sitting and standing are more difficult to differentiate. We think it is because these activities are similar in the view of sensors.
- Standing has a high false negative, might due to lack of training data samples.
- Typing and Standing, Walking and Standing, Typing and Laying are harder to be classified from each other. This might be because of multiple activities are taken as the same time, but we only support user to tag one activity per data sample.

$$\hat{f}^{bag}(x) = \frac{1}{B}\sum_{b=1}^{B}\hat{f}^{(b)}(x) \tag{1}$$





Future work

From this project, we noticed the following facts: 1) Adding more data is still improving our test error on the margin. We would like to try to add more data to see how much this will improve the model. 2) Having a better definition of activities would improve this model. We would like to allow the possibility of tagging multiple activities at the same data point, since some of activities can be done simultaneously. 3) Sensor data collection from different types of phones can be significantly different. The reason for this could be the sensors calibration and precision are different for each phone. We end up using the data from the same phone. 4) There exists an activity pattern for each user. Using one user's training data to predict another user's behavior is performance worse than predict this user's behavior. Personalized model for each user might improve the prediction accuracy. We would like to solve these problem in the future.

References and Acknowledgements

Thanks for Professor Andrew Ng, Professor John Duchi, TA Rishabh Bhagava's instruction.

- Shu Chen and Yan Huang, "Recognizing human activities from multi-modal sensors," Intelligence and Security Informatics, 2009. ISI '09. IEEE International Conference on, Dallas, TX, 2009, pp. 220-222.
- Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine, Davide Anguita, et. al.
- Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. A Public Domain Dataset for Human Activity Recognition Using Smartphones. 21th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013. Bruges, Belgium 24-26 April 2013.gnal Detection and Estimation. New York: Springer-Verlag, 1985, ch. 4.