AN ABSTRACT OF THE THESIS OF

Michael M. Anderson for the degree of Master of Science in Computer Science

presented on June 7, 2013.

Title: Activity Detection on Free-Living Data Using Change-Point Detection

Abstract approved: ______________________________

Weng-Keen Wong

Activity detection on time series sensor data is a rapidly emerging field with a lot of potential real-world application. In previous work researchers have unrealistically assumed that they know at which ticks in a time series the given subject stops doing one activity and starts doing another. In this thesis we explore the feasibility of segmenting realistic free-living time series data using techniques from change-point detection, and then classifying the predicted segments using standard supervised learning techniques. We compare this to the popular approach of splitting the time series into small windows of fixed-length, predicting on each window with a classifier, and then smoothing over the predictions with an HMM. We find that the HMM approach clearly outperforms the change-point detection approach, but that change-point detection may be promising given a modeling assumption that is appropriate to the data.

Activity Detection on Free-Living Data Using Change-Point Detection

by

Michael M. Anderson

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented June 7, 2013
Commencement June 2013

Master of Science thesis of Michael M. Anderson presented on June 7, 2013.

APPROVED:

Major Professor, representing Computer Science

Director of the School of Electrical Engineering and Computer Science

---

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Michael M. Anderson, Author

ACKNOWLEDGEMENTS

I would like to acknowledge and thank my advisor, Dr. Weng-Keen Wong, for providing me with a clear and overarching vision of this project at all stages of its development, as well as his general expertise and helpfulness with all of the rough patches, sticking points, and unexpected problems that invariably accompany an endeavor of this magnitude. I would also like to thank the other members of my committee, Dr. Alan Fern, Dr. Raviv Raich, and Dr. Hector Vergara, for giving their time to help supervise the final stages of this project. Thanks go to all of my family and friends who encouraged and stood by me along the way. They are too numerous to name but they know who they are.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1: Results

1.1 Change-Point Detection

Results for our change-point detection experiments are given in Figures 1.1-1.3. We hypothesized that the performance of the change-point detection algorithms would depend heavily on the threshold level for change prediction. This was tested by varying the average number of times per second that the algorithms falsely predicted a change. A large number of such false positive rates per second were tested, but for the sake of brevity only a representative sample of $\{0.005, 0.01, 0.05, 0.1\}$ are shown here.

In the OSU Hip experiments, control charts outperformed KLIEP in terms of detection time (Figures 4.1.2, 4.1.4, 4.1.6), while the accuracy results were mixed. It is generally expected that the accuracy curve as a function of the false positive rate will be unimodal: very large windows extend into multiple activities and confuse a classifier, while very small windows do not contain enough information to be discriminative. This unimodal behavior is shown in the control chart results, but not in the KLIEP results (Figures 4.1.1, 4.1.3, 4.1.5). Follow-up experiments showed that the peak in KLIEP accuracy performance occurred between false positive rates of 0.2 and 0.3 for each of the three classifiers.

Further investigation indicated that across the OSU Hip dataset the KLIEP algorithm was unable to detect many different activity changes without a very low score threshold value (and consequently very high false positive rates). Some qualitative plotting of the OSU Hip data showed that most of its activities have accelerometer amplitude values that strongly resemble draws from a multivariate normal distribution. Since control charts assume that the data is drawn from a distribution that is a member of that family, it is logical that control charts would outperform algorithms with different modeling assumptions on OSU Hip.

In the LiME experiments, KLIEP outperformed control charts in terms of accuracy across the board, and control charts outperformed KLIEP in terms of detection time across the board. (TODO: Need to interpret this, but not sure how.)

Finally, in a few cases (Figures 4.1.2, 4.1.6, 4.2.6) the detection time did not decrease as the false positive rate increased. On the face of it this would seem to be a non-sequitur, but this only happened in cases when accuracy also decreased (Figures 4.1.1, 4.1.5, 4.2.5). Smaller window sizes tend to be correlated with decreased detection times, but it is possible that predicting on smaller windows, if they happen to be less discriminative, can actually increase the time required for the classifier to start correctly predicting the ground-truth activity. Additionally, the given increases in detection time were small and near standard error.

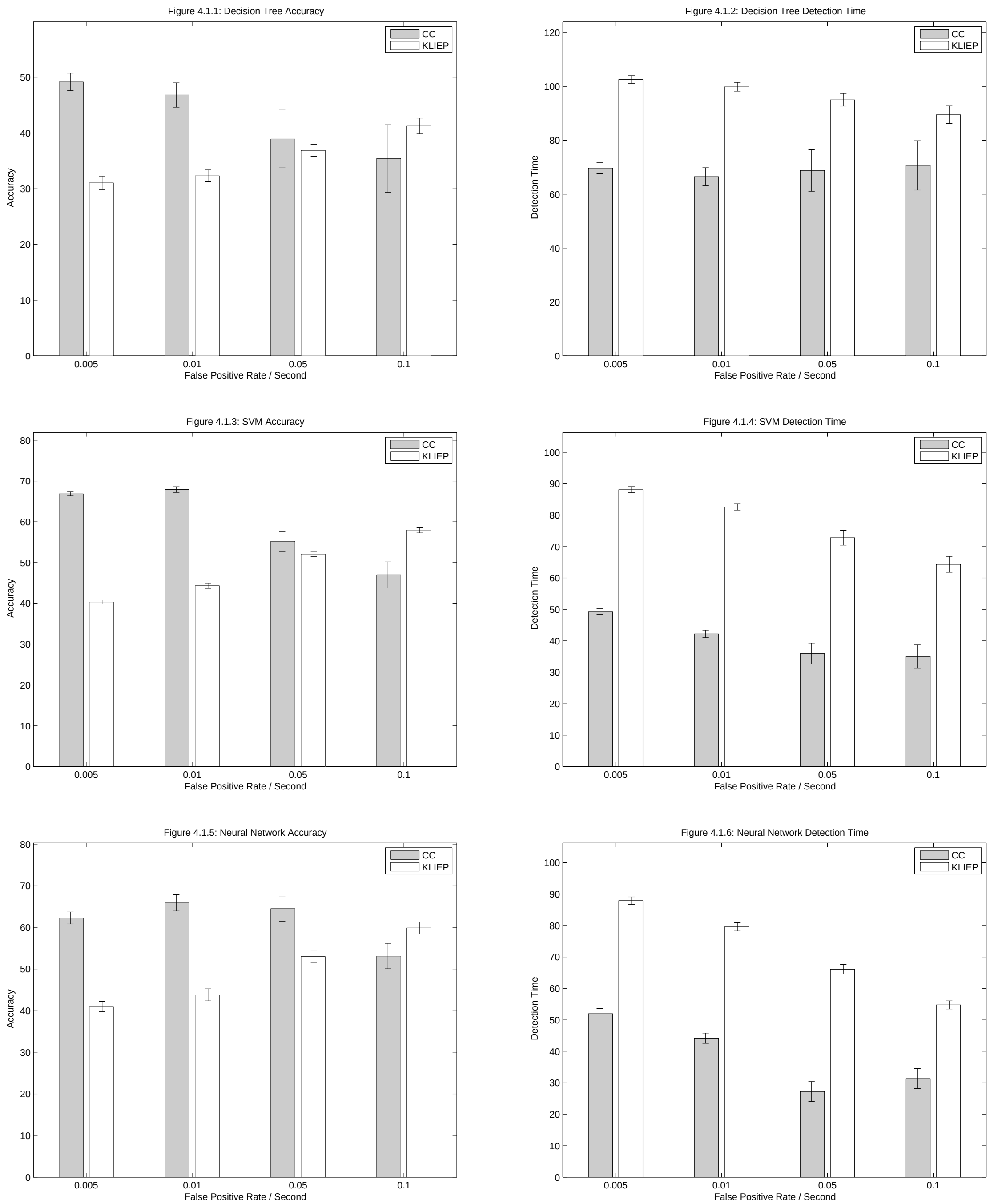

Figure 1.1: OSU Hip Results. Graphs are organized into rows by base classifier, and columns by evaluation metric. Change-point detection result were averaged over 30 splits into training, testing, and validation datasets, along with bars showing one standard error.

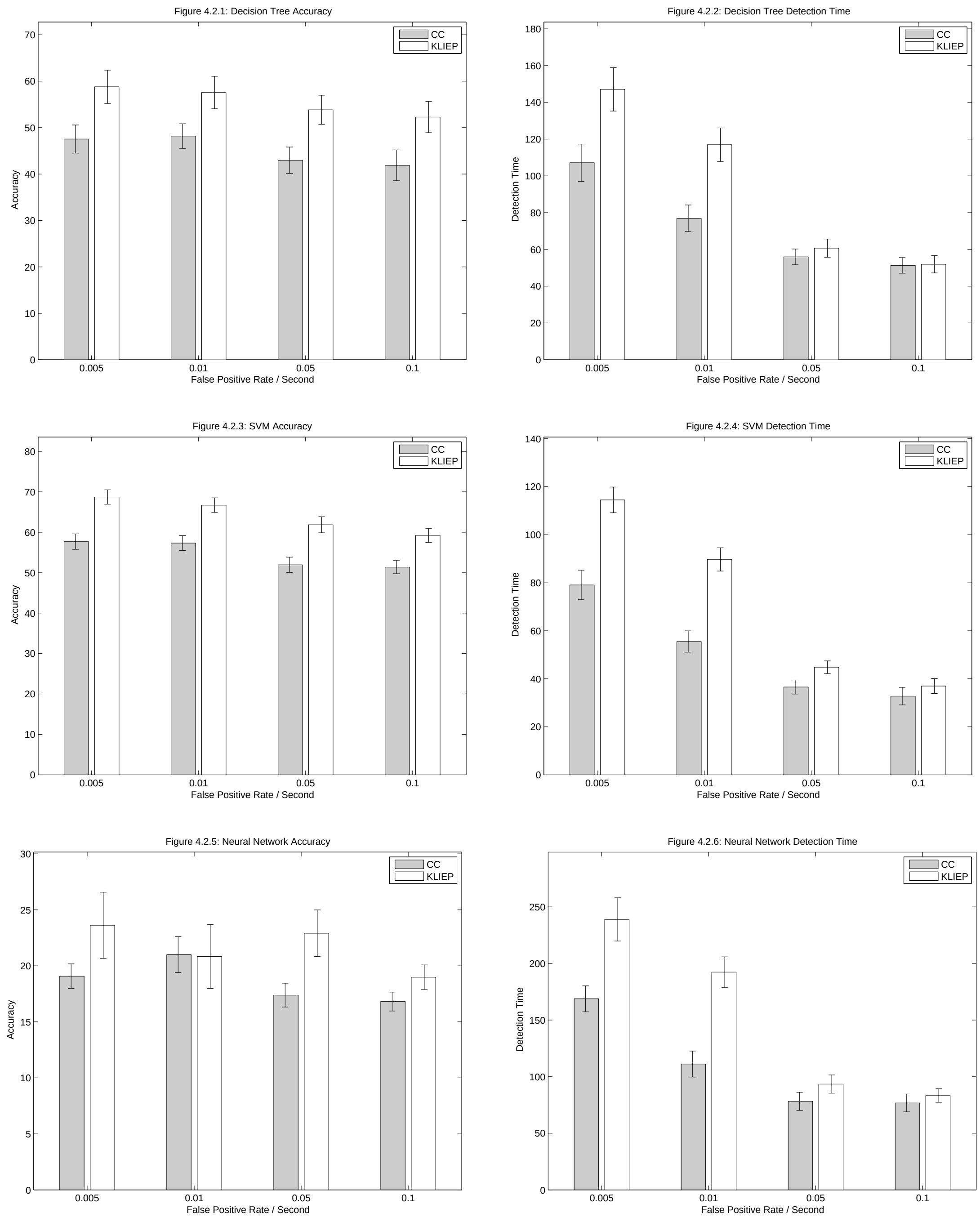

Figure 1.2: LiME Day 1 Results

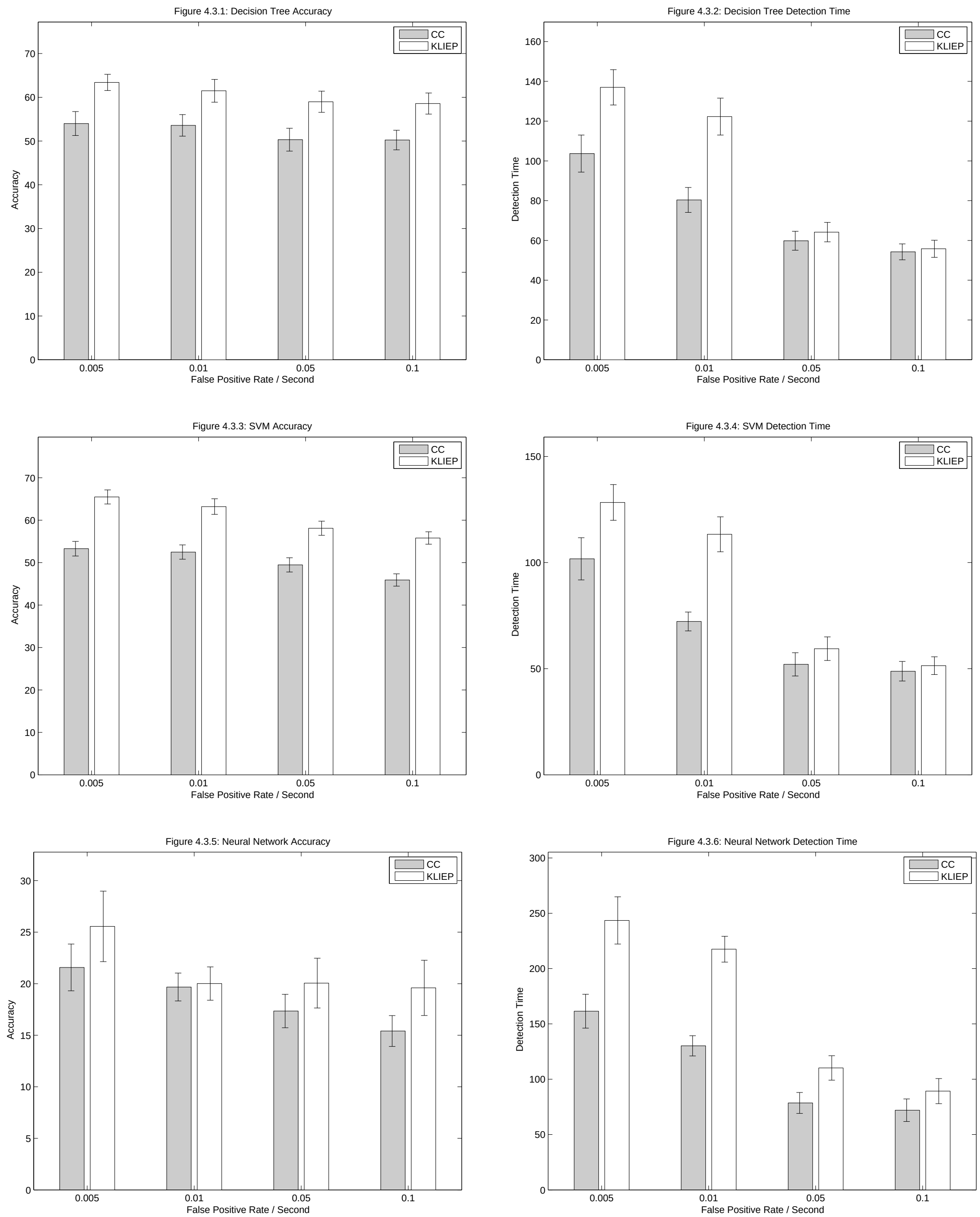

Figure 1.3: LiME Day 2 Results

1.2 HMM

Results for our HMM experiments are given in Figures 1.4-1.6. Each HMM experiment was performed by splitting time series into windows of fixed length for featurization, and results for windows of length $\{10, 12, 14, 16, 18, 20\}$ seconds are shown.

For both the SVM and decision tree classifiers, accuracy and detection time was strong across all three datasets, and also stable with respect to window size. Further experiments on the OSU Hip dataset showed that the HMM when paired with these classifiers tends to be stable with window sizes that are greater than a few seconds, which seems to be the amount of time required to be informative. Neural networks performed somewhat more poorly and erratically across the board.

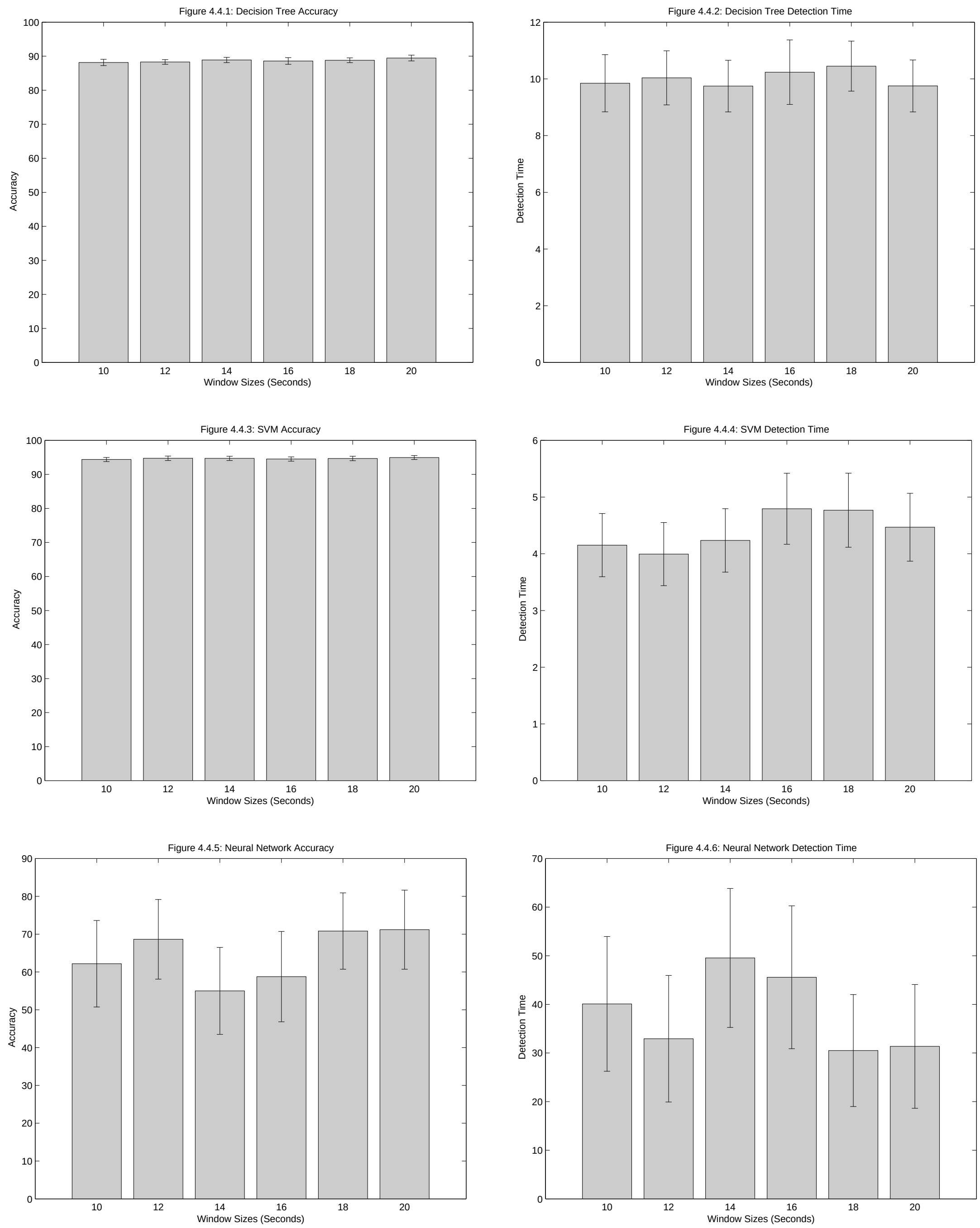

Figure 1.4: OSU Hip HMM Results. Graphs are organized into rows by base classifier, and columns by evaluation metric. HMM results were averaged over 10 splits into training (base classifier), validation, training (HMM), and testing datasets, along with bars showing one standard error.

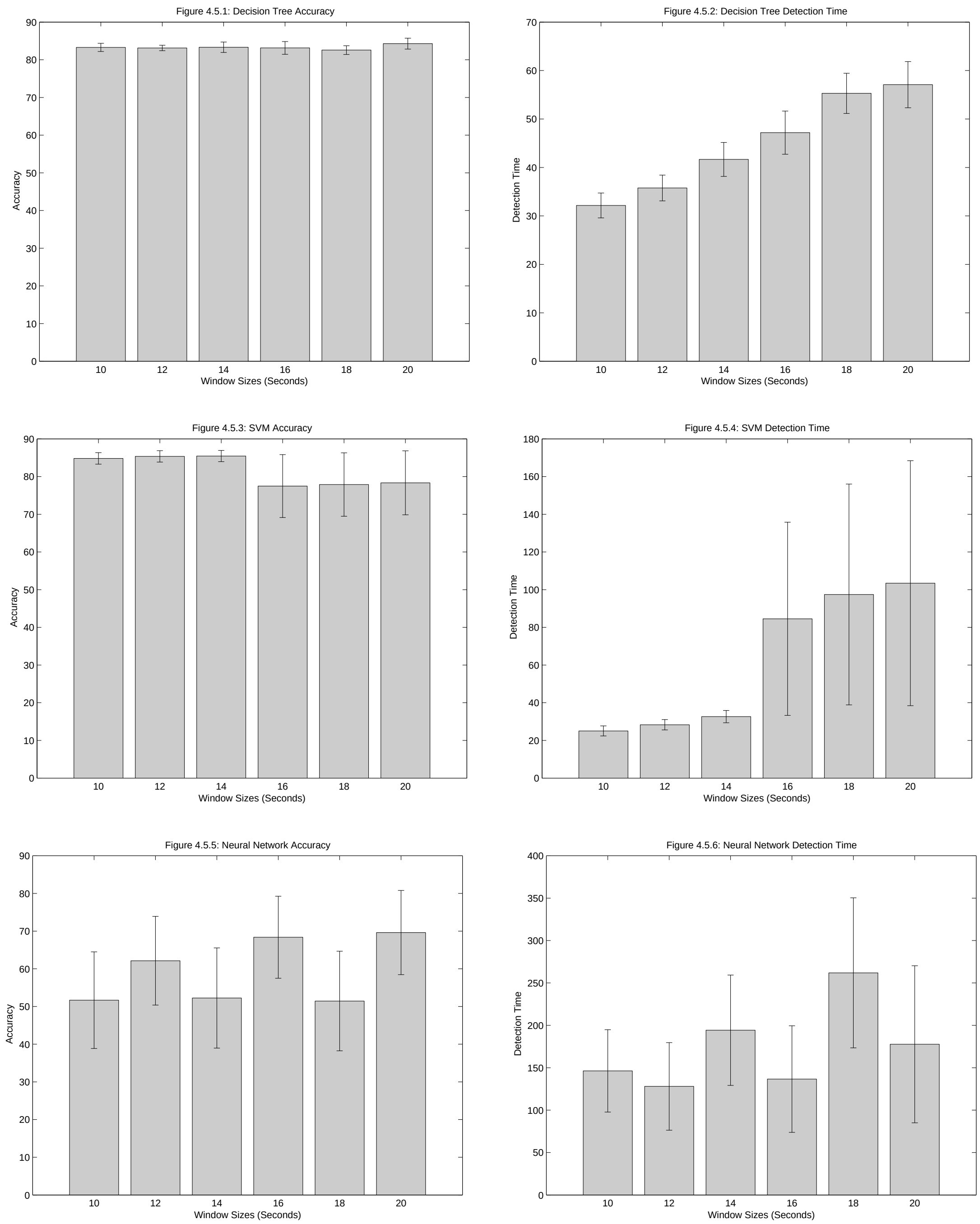

Figure 1.5: LiME Day 1 HMM Results

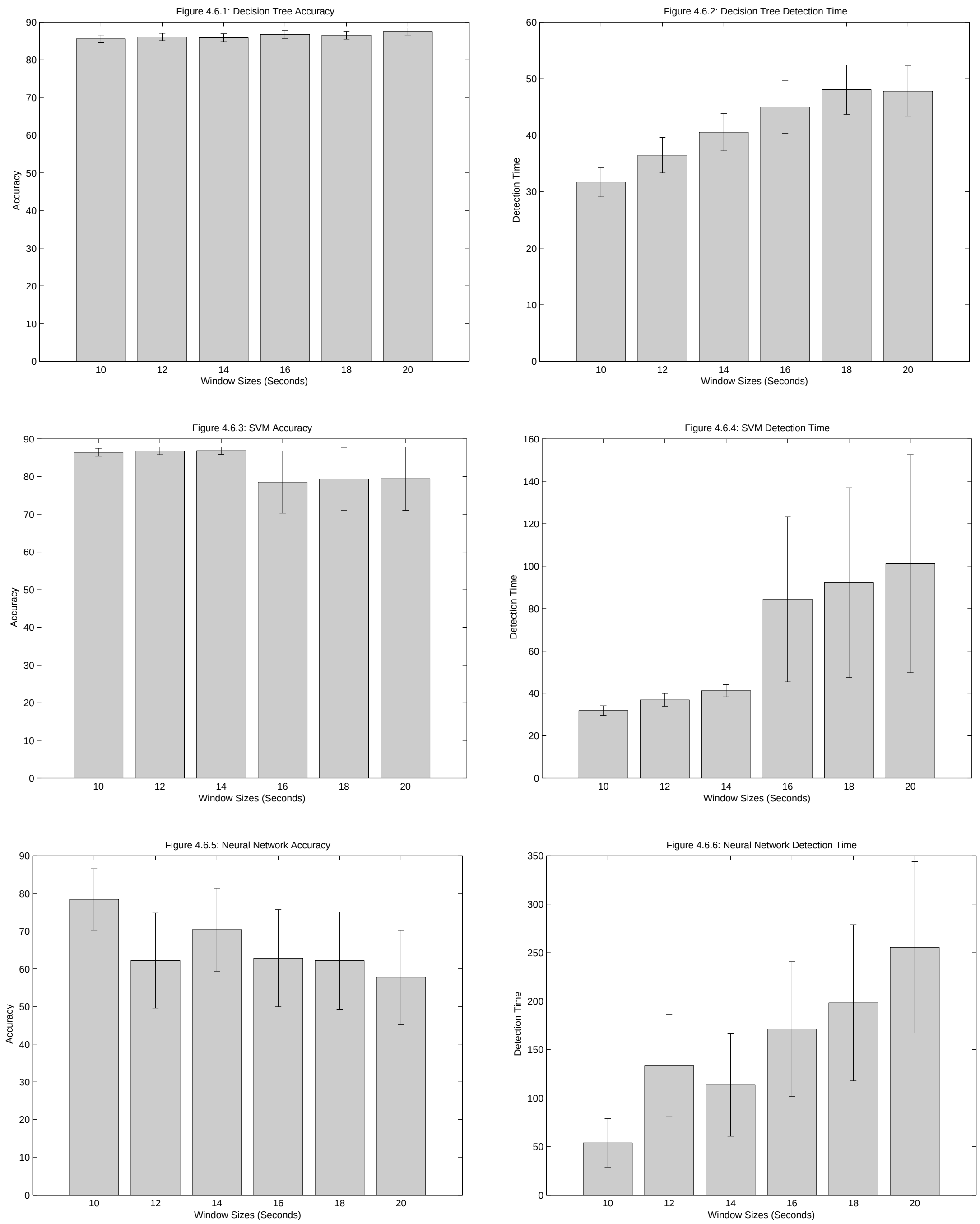

Figure 1.6: LiME Day 2 HMM Results

1.3 Discussion

Our results clearly show that the HMM approach outperformed the change-point detection approach, both in terms of accuracy and detection time, regardless of the dataset and base classifier. This was expected, as our change-point detection algorithms did not perform expecially well at predicting changes at the correct locations, and also because HMMs are a well-established and well- grounded approach in sequential, time-oriented domains.

A further point of interest was that SVM clearly beat out the other two base classifiers, and that the faster and simpler decision tree model did fairly well against neural networks. This result is significant because much of the previous research that has formulated activity detection as a supervised learning problem has used neural networks exclusively.

Bibliography