

## **Figure 1. Percentage Misclassified By Activity Performed**

In an experiment volunteer participants were asked to wear smartphones that would collect sensor data while they performed one of six activities: Laying down, sitting, standing, walking, walking down, or walking up.

Two classifier models, Multinomial Logistic Regression and Random Forest, were trained on a set of samples and then used to predict which activity the participant was performing based on the sensor data from an unseen test set.

This figure focuses on the issue of how well each model was able to predict each of the six activities by comparing the percentage of misclassifications each model produced on the test set for each activity. Focusing on this issue helps to further the main analysis of comparing the two classifier models by examining the performance of each in greater detail.

Multinomial Logistic Regression more often correctly predicted the activities of sitting, standing, and walking down. The model's error rates for each of these activities, rounded to the nearest percent, were: Sitting 11%, Standing 3% and Walking Down 2%.

Random Forest more often correctly predicted the activities of walking and walking up. The model's error rates were: Walking 0% and Walking Up 0.5%.

The models performed equally well on laying down, with both predicting this activity with 100% accuracy.

The activity with the highest error rate for both models was sitting, with Multinomial Logistic Regression predicting incorrectly on 11% of test samples and Random Forest predicting incorrectly on 14%.