Preprocessing:

* Data was separated by subject and further separated by day and further separated by types of information(orientation, gps, acceleration, environment, activity, posture)
* Each record was recorded less than a second apart, thus over a million records(Excel could only view that many at a time)
* Several attributes overlap different datasets(ex: environment attributes included in gps data) and have “UNKNOWN” values

Post-processing:

* Only used attributes directly describing subject motion(x, y, z, pitch, roll, azimuth)
* Cut out “useless” data(gps and activity datasets and any empty or attributes with unknown fields)
* Pared data down to one record per second
* Combined useful attributes into single data file
* Used posture dataset as labels