A Machine Learning Approach to Sleep Quality Prediction for Inpatient Rehabilitation

Allison Fellger

GONZAGA UNIVERSITY

ALLISON FELLGER, ALEXA ANDREWS, GINA SPRINT, DOUGLAS WEEKS, ELENA CROOKS | GONZAGA UNIVERSITY | SCHOOL OF ENGINEERING AND APPLIED SCIENCES

OVERVIEW

ACTIGRAPHS are wearable sensors used to collect activity and sleep time series data. Our approach utilizes similarities among historical sequences of data to train machine learning algorithms to predict nighttime sleep duration for STROKE and TRAUMATIC BRAIN INJURY (TBI) patients during rehabilitation. Our approach is suitable for point-of-care and remote monitoring to detect changes in sleep for individuals recovering from stroke and TBIs.

DATA COLLECTION

- MOTIONLOGGER Actigraph watches, by Ambulatory Monitoring, Inc
- 17 TBI and stroke patients
- 7+ days of data per patient
- MINUTE-BY-MINUTE activity data collected 24 HOURS per day

DAYTIME (DT) FEATURES	NIGHTTIME (NT) FEATURES	
DAYTIME ACTIVITY RATIO (DAR)	TOTAL SLEEP TIME (TST)	LONGEST SLEEP BOUT
MEAN ACTIVITY COUNT	SLEEP EFFICIENCY	MEAN SLEEP BOUT
MEDIAN ACTIVITY COUNT	SLEEP ONSET LATENCY	SLEEP TRANSITIONS
ACTIVITY COUNT STD DEVIATION	WAKE AFTER SLEEP ONSET	AROUSAL INDEX

PREDICTION MODEL

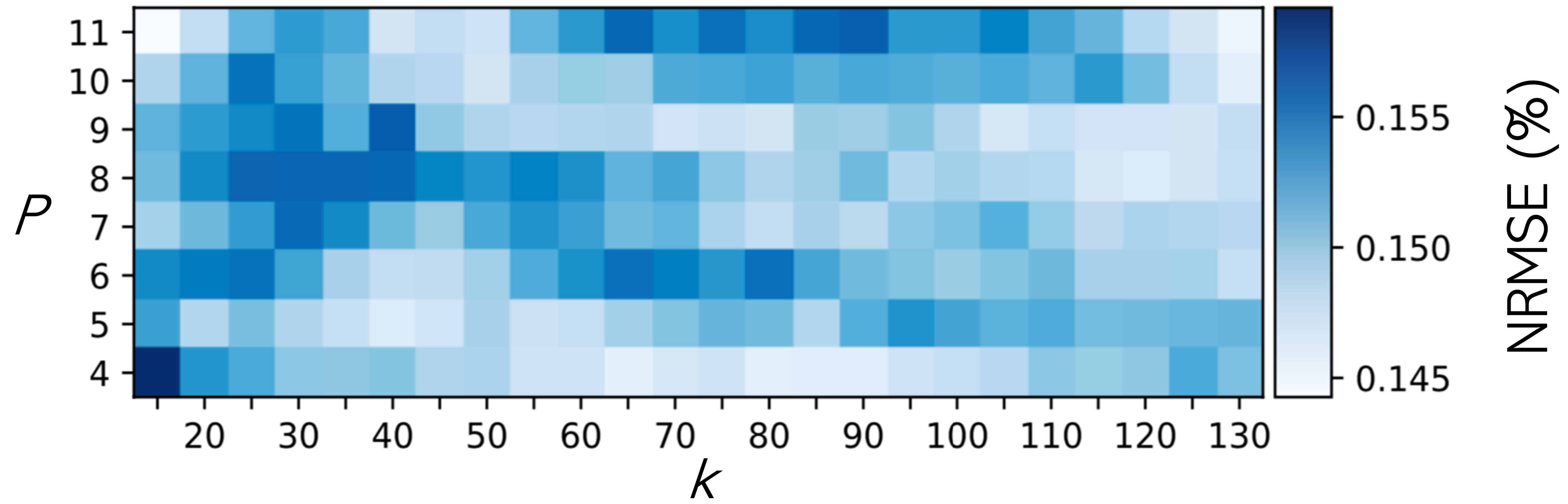
- *k*-NEAREST NEIGHBORS (KNN) model was used to identify *k* similar Actigraphy sequences based on DT and NT features during preceding periods
- RANDOM FOREST REGRESSION was used to predict the total nighttime sleep duration based on similar sequences
- *P* denotes the number of preceding periods used to formulate predictions

$DT_0 >> NT_1 >> DT_2 >> NT_3 >> DT_4 >> NT_5 >> DT_6 >> NT_x ...$

NT_3 NT_5 $NT_{x...}$

Example sequencing for P = 3

RESULTS



TOP NT TOTAL SLEEP TIME PREDICTION RESULTS

MODEL	P	K	NRMSE
RF_{LOOCV}	11	[all sequences]	14.40%
RF_{KNN}	11	15	14.43%
RF_{KNN}	11	130	14.51%
RF_{KNN}	10	130	14.56%
RF_{KNN}	11	40	14.69%

CONCLUSIONS

By tuning parameters related to our regression algorithm, we obtained a NORMALIZED ROOT MEAN SQUARE ERROR of 14.40%. Through experimentation, WE FOUND THE OPTIMAL *P* VALUE TO BE 11 PERIODS, which is the maximum value of *P* for this dataset. To elaborate, this means that each training sample consisted of the features from six DT periods (DT_0, DT_2, DT_4, DT_6, DT_8, DT_{10}) and five NT periods (NT_1, NT_3, NT_5, NT_7, NT_9), which were used to predict TST during NT_{11}. The poorest performing P values included P={1, 2, 3}. These were excluded from the heat map above to elucidate the patterns of the larger *P* values. These results suggest LONGER SAMPLES CAN IMPROVE ACCURACY. They also demonstrate that similar results can be obtained with lower *P* values if *k* is increased to compensate.