# Deep Sequence Learning on Mobile Data for Mood Prediction

Dimitris Spathis[1], Sandra Servia-Rodriguez[1], Katayoun Farrahi[2], Cecilia Mascolo[1], Jason Rentfrow[1]
ds806@cam.ac.uk, ss2138@cam.ac.uk, kf2u17@ecs.soton.ac.uk, cm542@cam.ac.uk, pjr39@cam.ac.uk
[1]University of Cambridge, [2]University of Southampton

Mental health illnesses are among the most prevalent diseases worldwide, but only a small part of the population receives treatment. Psychologists have been measuring mental health for decades using pen and paper surveys and controlled laboratory experiments. The pervasiveness of smartphones and wearables has enabled a near real-time detection of clinical outcomes. Self-reported mood is associated with clinical outcomes like depression, longevity and reduced mortality risk. However, most of the studies that investigate how mood reports collected using smartphones can help to improve mental health, have been conducted through controlled experiments with limited number of participants (LiKamWa et al., 2013). Common tasks include prediction of mood, stress, depression, sleep quality or academic performance, using smartphone sensors and appropriate apps.

Deep neural networks, have become increasingly useful in modeling high-dimensional, non-linear data such as time-series. Unlike other time-series methods like Hidden Markov, or autoregressive models, they can learn patterns from multiple time-series and therefore different users, without resorting to concatenation tricks. On the other hand, regression algorithms like random forests cannot handle sequence data and require some kind of laborious feature extraction. Recurrent Neural Networks can deal with these issues but they operate by mapping an input sequence to an output sequence of the same length, which is not very flexible for forecasting. However, recent encoder-decoder models overcome this limitation by converting the source sequence into a *thought* vector which is passed through a decoder to produce a forecast (Sutskever et al., 2014).

In this study, we propose a multi-task encoder-decoder recurrent neural network to accurately forecast the mood horizon from previous self-reported mood collected using mobile devices. It involves two RNN networks, the *Encoder* that maps the past mood to a fixed length vector with the size of the output, and the *Decoder* that reconstructs this vector to arbitrary-length future steps. Unlike most previous work, we do not treat mood as a binary state but we forecast precise score ranges, while the model operates directly on the raw, noisy, sparse time-series.

We evaluated our approach on the Emotionsense dataset, the longest-running experiment in mobile sensing and mood self-reporting, tracking more than 17.000 users for over 3 years. For our experiments we used a subset that includes more than $500$ users and $170,000$ self-reports. The reports come from the affect grid, a 2D grid of valence-arousal ranging [0,1]. We determine how many previous reports are required to accurately predict future mood, showing that 3 weeks of occasional reporting is the optimal period; going further back is not informative. The accumulation of errors as we forecast further into the future is notable. We exploit the fact that the two affect dimensions –valence and arousal– are related, to perform multi-task learning (Taylor et al., 2017). The multi-task setup which forecasts two sequences at the same time, achieves better results than training separate models or other baselines. Our model is able to forecast tomorrow's mood better than feature-based classifiers or just repeating the past, achieving $\pm 0.14$ error on the affect grid for tomorrow and $\pm 0.16$ for 7 days later.

A visualization of the model's internal learned *black-box* representations with dimensionality reduction revealed the latent space of mood forecast; past mood is getting clustered with respect to the mood of the next day. Inspecting the behavior of individual neurons of the Decoder, we observe that different neurons learn different non-linear sequential patterns of future mood. Given that human mood might vary a lot within a single user, and especially across a population, we studied how well the model performs for users and groups with different self-reported mood variability and distinct personality traits. We found that the performance increases with higher emotional stability and openness of the users. Increased future mood variability is correlated with higher model error.

Our experiments showed the efficacy of an encoder-decoder model to forecast mood based on past self-reports. However, this requires occasional user involvement and self-reporting that might introduce bias and sparseness. What if we could do it without bothering the user? Our ongoing work on tomorrow's mood prediction based only on passive mobile sensors and user metadata on a larger sample, shows promising results. For this new task, we propose a multi-modal network with recurrent-convolutional parts for the sensors (accelerometer/microphone), and dense layers for the user surveys (demographics, personality traits), and concatenate them with merge layers. We report accuracy of $\sim 66\%$ (disjoint user test set) on predicting the binary dimension of valence on the affect grid. Ablation studies show that the static personality features contribute more, while sensors slightly boost the performance. By considering only on the extreme happiness/sadness, accuracy reaches $\sim 80\%$, hinting that the neutral mood users introduce substantial noise.

## References

LiKamWa, Robert; Liu, Yunxin; Lane, Nicholas D., and Zhong, Lin. Moodscope: Building a mood sensor from smartphone usage patterns. In *MobiSys '13*. ACM, 2013.

Sutskever, Ilya; Vinyals, Oriol, and Le, Quoc V. Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*, pages 3104–3112, 2014.

Taylor, Sara Ann; Jaques, Natasha; Nosakhare, Ehimwenma; Sano, Akane, and Picard, Rosalind. Personalized multitask learning for predicting tomorrow's mood, stress, and health. *IEEE Transactions on Affective Computing*, 2017.