

Development of Conversational AI for Sleep Coaching Programme

Heereen Shim

KU Leuven, Campus Group T, eMedia Research Lab, Leuven, Belgium
KU Leuven, Department of Electrical Engineering (ESAT), STADIUS, Leuven, Belgium
heereen.shim@kuleuven.be

Abstract

Almost 30% of the adult population in the world is experiencing or has experience insomnia. Cognitive Behaviour Therapy for insomnia (CBT-I) is one of the most effective treatment, but it has limitations on accessibility and availability. Utilising technology is one of the possible solutions, but existing methods neglect conversational aspects, which plays a critical role in sleep therapy. To address this issue, we propose a PhD project exploring potentials of developing conversational artificial intelligence (AI) for a sleep coaching programme, which is motivated by CBT-I treatment. This PhD project aims to develop natural language processing (NLP) algorithms to allow the system to interact naturally with a user and provide automated analytic system to support human experts. In this paper, we introduce research questions lying under three phases of the sleep coaching programme: triaging, monitoring the progress, and providing coaching. We expect this research project's outcomes could contribute to the research domains of NLP and AI but also the healthcare field by providing a more accessible and affordable sleep treatment solution and an automated analytic system to lessen the burden of human experts.

1 Introduction

Insomnia is one of the most common sleep disorders with a high prevalence. Approximately one-third of adults experience one or more of the symptoms of insomnia (Roth, 2007). The consequences of insomnia include not only individual problems but also societal issues, such as daytime fatigue, low energy level, which can cause depression, and even increased risk of accidents (Leger et al., 2014). The cost associated with insomnia, including direct and indirect costs, in the US is around 92.5 to 107.5 billion USD per year (Stoller, 1994).

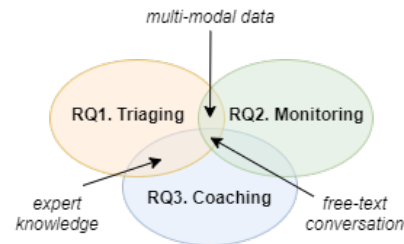


Figure 1: Overview of the proposed PhD research project: Three research questions (RQs) and the overlapped research topics.

Cognitive behaviour therapy for insomnia (CBT-I) is one of the most effective solutions to treat insomnia. During CBT-I process, a person who is suffering from symptoms of insomnia (patient) will consult with a CBT-I provider (therapist) who provides support to identify behaviours, thoughts, and feelings that are related to the symptoms. Since its goal is to change the potential causes, both behavioural and cognitive factors, it produces long-lasting improvement in the condition of insomnia, compared to the medication treatment (Morin et al., 2006).

Despite its effectiveness, CBT-I treatment has a limitation. From a patient's perspective, the treatment cost is high and, from clinician's perspectives, the number of patients that potentially can be treated is significantly large (Edinger and Means, 2005). Consequently, many researchers and engineers have been working on developing more accessible and affordable therapy solutions that can also lessen the burden of clinicians, such as internet- or mobile-based computerised therapy tools (Ström et al., 2004; Ritterband et al., 2009; Vincent and Lewycky, 2009; Lancee et al., 2012). These studies explored opportunities for applying technologies to automate treatment process. However, the conversational aspect, which is the core of the in-person treatment, has been neglected.

In this study, we will explore the possibilities of developing conversational AI (artificial intelligence) to make a computerised sleep therapy tool to be more close to in-person therapy. Since this research field is still in its infancy, we consider a sleep coaching programme targeting healthy people who would like to optimise their sleep, rather than a sleep therapy for patients with the chronic sleep disorder. Also, the goal of this study is to provide a user-friendly interface for users and to support human experts, rather than to replace or eliminate human-in-the-loop. Therefore, we mainly focus on two things: 1) adding a conversational feature that allows users to provide inputs to enable a natural conversation between human and the system; and 2) adding an analytic feature that automates processing user inputs to support decision making. The main research questions of this project lie under three processes of the sleep coaching programme: triage, monitor the progress, and provide coaching. Overview of research questions and overlapped components are illustrated in Figure 1.

We start our research by analysing in-person sleep treatment and existing automated tools and identifying missing gaps to decide research questions (Section 2). Then we revisit the research questions and explain the methodology to address each question (Section 3). As the first step, we introduce a pilot study and present a preliminary result to discuss and provide next steps (Section 4). Finally, we conclude this paper by summarising research questions and research plan of this PhD project (Section 5).

2 Related Work

In this section, we will first provide a brief overview of sleep treatment of CBT-I made by a human expert. Secondly, we will review the existing methods of automated sleep treatment tools. Lastly, we will identify missing gaps and introduce research questions.

2.1 In-person treatment

CBT-I is a sleep treatment that focuses on investigating the relationship between how we behave, how we think, and how we sleep. To achieve this, the treatment consists of multiple components: stimulus control, sleep restriction, sleep hygiene education, relaxation training, and cognitive restructuring (Perlis et al., 2006; Morin and Espie, 2007; Belanger et al., 2006). Stimulus control aims

to change associations between the bedroom with habits that make sleeping more difficult (Bootzin and Perlis, 2011). Sleep restriction treatment requires patients to limit time spent in bed in order to resolve the mismatch between the time in bed and sleep time (Spielman et al., 2011). Sleep hygiene focuses on educating the patient to avoid behaviours that influence sleep (Kleitman, 1987; Hauri, 1991). Relaxation training is given to help reduce the racing thoughts (Ong et al., 2014). Cognitive restructuring targets to break the vicious circle between inaccurate thoughts about sleep and behaviours that contribute to insomnia. Standard CBT-I treatment includes three or more of these components.

Conversation between patient and therapist plays a critical role in in-person sleep treatment. In the first session of treatment, the patient will provide complaints of their sleep and the therapist will determine whether the patient is appropriate for CBT-I treatment. To identify this, the therapist should assess the patient based on the clinical interview and the completed questionnaires. Once it is determined that the patient is appropriated for the treatment, the therapist will select the treatment components and structure plan tailored to the patient. The remaining sessions will be followed depending on the stage of treatment and the degree of patient compliance. Therefore, it is important that the therapist monitors the progress, identifies the patient's difficulties, provides personalised support, and encourages the patient to complete the treatment.

2.2 Computerised treatment

One of the earliest approaches is a research work by Ström et al. (2004) investigating the feasibility of an internet-based CBT-I. They proposed a self-help program that patients provide their information and progress by completing questions and questionnaires via the internet. However, their method does not provide automated analytic features to support monitoring process done by human experts.

Later on, Ritterband et al. (2009) proposed a fully automated sleep treatment tool. They proposed an automated algorithm that can produce a personalised recommendation for sleep restriction. It also automatically sends emails of reminders. All intervention was delivered without human support and the outcome was comparable to in-person treatment. Nevertheless, it still misses the interactive

conversational feature that participants could report their specific concerns or difficulties of complying the treatment.

Later studies also did not explore the opportunity to get feedback or support. For example, both Vincent and Lewycky (2009) and Lancee et al. (2012) developed online treatment methods for insomnia, but neither of them offered automated support feature where the participants could contact when facing issues or difficulties to follow the sessions. Until recently, conversational aspects have been neglected (Kuhn et al., 2016; Werner-Seidler et al., 2017; Horsch et al., 2017).

2.3 Missing gaps and research questions

So far, conversational aspects, which plays a critical role in in-person sleep treatment, are less studied. Our main hypothesis is that conversational AI will enable natural interaction between users and a system, throughout the treatment process, to make a computerised treatment be more close to in-person treatment. To this end, we will focus on the following research questions:

RQ1. How to triage via conversation together with the completed questionnaires?

RQ2. How to monitor the progress during the sleep coaching programme?

RQ3. How to understand a user-specific situation for personalised coaching programme?

3 Research Plan

RQ1: How to triage via conversation together with the completed questionnaires?

Answer to this research question entails three sub-tasks: The first sub-task is to assess users' complaints and identify sleep-related issues and its impacts to identify potential causes. The second sub-task is to ask follow-up questions to clarify ambiguous statements and differentiate causes that have similar impacts. The third sub-task is to explain the assessment result.

Complaints assessment

Assessing the users' complaints can be reformulated as a classification task. One of the most similar approaches is a recent study conducted by Shim et al. (2020). They used neural networks-based multi-label classifier to detect pre-defined sleep issues from free-text. What makes our study more challenging is that 1) we aim to assess not

only sleep issues but also impacts to identify underlying causes. And 2) we aim to incorporate the completed questionnaire results, which is different modality from free-text. For the first challenge, a naive approach is to build three separate classifiers for sleep issues, impact, and causes. It is limited, however, because these three entities are connected, such as there are causal links between sleep issues to impacts. Therefore, as the first step, we will build a directed graph, such as a Bayesian Network (BN) that each node represents the observed results of each entity. Since each node can be either free-text classification results or questionnaire results, we can address the second challenge, too. We will also explain other benefits of implementing the BN in the following paragraphs.

Follow-up question

In this project, we will explore the potential of conversational environment that the system can interact with patients. One of the benefits is that the system can actively search for additional information when it is needed. For example, the free-text inputs from users can be ambiguous. Also, multiple sleep issues could result in similar impacts so that it requires further assessment that differentiates between two or more conditions. We hypothesise that asking follow-up questions will solve these challenges by clarifying and refining it. Then the real challenge becomes how to decide 'when to ask?' and 'what to ask?'. A study by Middleton et al. (2016) addressed these challenges by framing a triage as a sequence of questions and answers. To achieve this, they encoded expert knowledge into a graph structure; each possible questions is linked to the possible answers, each of which is linked to a follow-up question. Since this approach requires human resource, we foresee to work with experts in sleep domain to encode the domain knowledge into structured form, such as a graph. Also, we will investigate whether the BN can select the most appropriate follow-up question. We will describe a preliminary experimental result of this approach in Section 4.

Triage result

The end task of the first research question is to support triage via text-based conversation and by explaining the assessment result: what was the main complaint from the user and which habits were associated with the detected complaints. We plan to take a similar approach with Chen et al. (2020)

who implemented BN on top of neural networks to provide interpretability. However, compared to their task, our task is more challenging because not all information is given from the beginning. Therefore, we will model uncertainty to deal with unknown information to triage. For evaluation, we will follow a recent study by [Razzaki et al. \(2018\)](#) and evaluate our triage system both quantitatively and qualitatively: Quantitatively, we will calculate precision and recall of detecting sleep issues and underlying causes. Qualitatively, we will rate triage flows with the help of an expert in sleep domain.

RQ2: How to monitor the progress during the sleep coaching programme?

One of the fundamentals of sleep treatment is to monitor the progress of a participant. To monitor the progress and analyse sleep patterns, we will use a sleep diary that is a widely used method to access people’s quality of sleep ([Monk et al., 1994](#); [Carney et al., 2012](#)). The traditional sleep diary consists of a log of sleep-related activities, including bedtime, wake up time, and sleep time and heavily relies on objective values, such as total time in bed (TIB), total sleep time (TST), and sleep efficiency (SE). However, to monitor the progress and understand the user-specific condition, subjective values and context should be also considered. To achieve this, we will use a narrative sleep diary written in free-text that contains not only the time information of sleep-related events but also the rich information of context. For example, a user can describe her/his sleep in free-text to explain not only how long they slept and how many times they woke up during the night but also the quality of sleep or feeling after sleep and what disturbed their sleep. Recent work by [Rick et al. \(2019\)](#) takes a similar approach to obtain qualifiable insights about the subjective experience of sleep by incorporating free-text user inputs. During this project, we will investigate combine different modalities, objective and subjective values, extracted from the narrative sleep diary to assess the progress.

RQ3: How to understand a user-specific situation for personalised coaching programme?

During the sleep coaching programme that helps user change their behaviour to improve their sleep, it is critical to provide personalised coaching programme tailored to a user. To achieve this, understanding the experience of a user and identifying

the user-specific issues and difficulties is essential. In this study, we will use free-text input from users that describe their experiences, thoughts, and feelings during the coaching programme. Specifically, we will consider a behaviour change programme and aim to build a model that performs aspect-based sentiment analysis on review comments from a user. Sentiment analysis (SA) is a widely used natural language processing (NLP) technique used to assess user experiences ([Liu, 2012](#)). Aspect-based sentiment analysis (ABSA) is a type of SA that aims to detect sentimental values expressed toward fine-grained aspects ([Pontiki et al., 2014](#)), rather than performing classification at the sentence level. Even though ABSA is widely studied, the majority of works are limited to the review of consumer products ([Do et al., 2019](#)). Recently, [Barahona et al. \(2018\)](#) conducted research on detecting mental health concepts for cognitive behaviour therapy from user inputs by reformulating it as sentiment analysis detecting negative sentiment. Similar to this, we will investigate using ABSA technique to detect concepts related to sleep health for providing personalised support and behaviour change programme by analysing user inputs during the sleep coaching programme.

4 Pilot study

We ran a pilot experiment to examine our assumptions for RQ1. The main goal of this pilot study as follows: 1) To build a model that classifies free-text user inputs. 2) To implement BN to select a follow-up question. Following subsections describes details of the experiments and results.

4.1 Dataset

Motivated by [Shim et al. \(2020\)](#), we collected free-text data via crowdsourcing platform. We also adapted their approach that asks participants *to imagine they are sitting at the doctor’s office and being ask to describe three different topics: sleep issues, the impact of their issues, and factors that might contribute to the issues*. We cleaned the data by dropping invalid input texts and annotated to create three datasets named issues, causes, and impact, respectively. [Table 1](#) summarises statistics of each dataset. More information about dataset can be found in [Appendix A](#).

| Dataset | #.class | Train | Test |
|---------|---------|--------|-------|
| Causes | 19 | 10,430 | 1,155 |
| Issues | 11 | 12,928 | 1,437 |
| Impact | 11 | 10,733 | 1,142 |

Table 1: Statistics of datasets. #.class refers the number of class categories. The numbers in Train and Test columns refer the number of data points.

4.2 Experimental settings

For classification, we used the pre-trained language model (Devlin et al., 2019) initialised with pre-trained weights and fine-tuned on our datasets. Implementation details are given in Appendix B. To evaluate the classifiers, we calculated macro-averaged precision, recall, and F1-score for issues, impacts, and causes, separately.

For selecting a follow-up question, we created a simple BN with three layers, which are sleep causes, issues, and impact. Details of the model are described in Appendix C. Since this is a pilot study, we only considered a few entities and created conditional probability tables (CPT) based on our limited knowledge. Note that the structure and CPT of BN are not clinically proved; The goal of this pilot study is to demonstrate the concept. At each iteration of question and answering, the BN updates its probability distribution at each node given information and selects the entity node with the highest probability of the entity is true. We qualitatively evaluated this approach.

4.3 Preliminary result and next steps

Table 2 summarises the classification result. The result shows that the model performs better on causes dataset than both on issues and impact datasets, even though there are more class categories in causes dataset. We conducted error analysis and observed that the trained models tend to misclassify similar classes. It implies that further assessment is needed to differentiate semantically close texts.

| Dataset | P (%) | R (%) | F-1 (%) |
|---------|-------|-------|---------|
| Causes | 94.9 | 91.7 | 93.2 |
| Issues | 87.4 | 79.9 | 82.7 |
| Impact | 79.9 | 72.3 | 75.2 |

Table 2: Classification results. All measures are macro-averaged per each class.

Figure 2 shows demonstrations of triage flow with BN. Each sub-figure shows a sequence of

follow-up questions and answers given condition: normal BMI¹ (2a) and high BMI (2b). It is worth noting that each flow selected different follow-up question after the classification model predicted the same results. It shows the possibility of using BN to select the most appropriate follow-up question given information.

Currently, we did not evaluate the system based on the final triage result and the appropriateness of follow-up question because our dataset contains only free-text describing sleep issues, causes, and impacts, separately. In our future study, we plan to follow a similar data collection protocol of Razzaki et al. (2018). They asked doctors to play patients based on given vignettes containing simple demographics, complaint, and other information that can be obtained by either open-ended or closed-ended questions.

5 Conclusion

In this paper, we propose a PhD project exploring potentials of developing conversational AI for a sleep coaching programme, which is motivated by CBT-I treatment, targeting healthy people who would like to optimise their sleep. The main goal of this PhD project is to develop NLP algorithms for conversational AI to allow the system to interact naturally with a user and provide automated analytic. To this end, we identified three research questions lying under three phases of the sleep coaching programme: triage, monitor, and support. We expect this research project’s outcomes could contribute to the research domains of NLP and AI but also the healthcare field by providing a more accessible and affordable sleep treatment solution and an automated analytic system to lessen the burden of human experts.

Acknowledgments

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 766139. This article reflects only the author’s view and the REA is not responsible for any use that may be made of the information it contains.

¹Body Mass Index (BMI) is a value derived from the mass (m) and height of a person (h), defined as m/h^2 .

```

***** LET'S START CHAT *****
Doctor: Okay, what is wrong with your sleep?
Type: I am always tired no matter how long I sleep.
=====
Detected impact:
energy (0.95%)
=====
Do you experience/have troubleFallingAsleep?
(confidence: 0.68)
Choose (y/n): n
Do you experience/have troubleStayingAsleep?
(confidence: 0.86)
Choose (y/n): y
Can you elaborate more?
Type: I wake up several times during night.
Doctor: Okay, I have enough information for triage.
Doctor: You main complaint is energy,
Doctor: because of troubleStayingAsleep.
Doctor: Potentially your habit_drinking is the cause.
Doctor: We will focus on changing this.

```

(a) When condition of 'high BMI' is given as *false*.

```

***** LET'S START CHAT *****
Doctor: Okay, what is wrong with your sleep?
Type: I am always tired no matter how long I sleep.
=====
Detected impact:
energy (0.95%)
=====
Do you experience/have snoring?
(confidence: 0.87)
Choose (y/n): y
Can you elaborate more?
Type: I snore aloud and it wakes me up.
Doctor: Okay, I have enough information for triage.
Doctor: You main complaint is energy,
Doctor: because of snoring.
Doctor: Potentially your high_BMI is the cause.
Doctor: We will focus on changing this.

```

(b) When condition of 'high BMI' is given as *true*.

Figure 2: Examples of triage flow with BN and neural networks-based text-classifier. Green coloured texts show user-inputs while white coloured texts show system outputs.

References

- Lina M Rojas Barahona, Bo-Hsiang Tseng, Yinpei Dai, Clare Mansfield, Osman Ramadan, Stefan Ultes, Michael Crawford, and Milica Gasic. 2018. Deep learning for language understanding of mental health concepts derived from cognitive behavioural therapy. In *Proceedings of the Ninth International Workshop on Health Text Mining and Information Analysis*, pages 44–54.
- Lynda Belanger, Josee Savard, and Charles M Morin. 2006. Clinical management of insomnia using cognitive therapy. *Behavioral sleep medicine*, 4(3):179–202.
- Richard R Bootzin and Michael L Perlis. 2011. Stimulus control therapy. In *Behavioral treatments for sleep disorders*, pages 21–30. Elsevier.
- Colleen E Carney, Daniel J Buysse, Sonia Ancoli-Israel, Jack D Edinger, Andrew D Krystal, Kenneth L Lichstein, and Charles M Morin. 2012. The consensus sleep diary: standardizing prospective sleep self-monitoring. *Sleep*, 35(2):287–302.
- Jun Chen, Xiaoya Dai, Quan Yuan, Chao Lu, and Haifeng Huang. 2020. Towards interpretable clinical diagnosis with bayesian network ensembles stacked on entity-aware cnns. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 3143–3153.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *NAACL-HLT (1)*.
- Hai Ha Do, PWC Prasad, Angelika Maag, and Abeer Alsadoon. 2019. Deep learning for aspect-based sentiment analysis: a comparative review. *Expert Systems with Applications*, 118:272–299.
- Jack D Edinger and Melanie K Means. 2005. Cognitive-behavioral therapy for primary insomnia. *Clinical psychology review*, 25(5):539–558.
- Peter J Hauri. 1991. Sleep hygiene, relaxation therapy, and cognitive interventions. In *Case studies in insomnia*, pages 65–84. Springer.
- Corine HG Horsch, Jaap Lancee, Fiemke Griffioen-Both, Sandor Spruit, Siska Fitrianie, Mark A Neerincx, Robbert Jan Beun, and Willem-Paul Brinkman. 2017. Mobile phone-delivered cognitive behavioral therapy for insomnia: a randomized waitlist controlled trial. *Journal of medical Internet research*, 19(4):e70.
- Nathaniel Kleitman. 1987. *Sleep and wakefulness*. University of Chicago Press.
- Eric Kuhn, Brandon J Weiss, Katherine L Taylor, Julia E Hoffman, Kelly M Ramsey, Rachel Manber, Philip Gehrman, Jill J Crowley, Josef I Ruzek, and Mickey Trockel. 2016. Cbt-i coach: a description and clinician perceptions of a mobile app for cognitive behavioral therapy for insomnia. *Journal of clinical sleep medicine*, 12(4):597–606.
- Jaap Lancee, Jan van den Bout, Annemieke van Straten, and Victor I Spoormaker. 2012. Internet-delivered or mailed self-help treatment for insomnia? a randomized waiting-list controlled trial. *Behaviour research and therapy*, 50(1):22–29.
- Damien Leger, Virginie Bayon, Maurice M Ohayon, Pierre Philip, Philippe Ement, Arnaud Metlaine, Mounir Chennaoui, and Brice Faraut. 2014. Insomnia and accidents: cross-sectional study (equinox)

- on sleep-related home, work and car accidents in 5293 subjects with insomnia from 10 countries. *Journal of sleep research*, 23(2):143–152.
- Bing Liu. 2012. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1):1–167.
- Katherine Middleton, Mobasher Butt, Nils Hammerla, Steven Hamblin, Karan Mehta, and Ali Parsa. 2016. Sorting out symptoms: design and evaluation of the ‘babylon check’ automated triage system. *arXiv preprint arXiv:1606.02041*.
- Timothy H Monk, CHARLES F REYNOLDS III, David J Kupfer, Daniel J Buysse, Patricia A Coble, Amy J Hayes, Mary Ann Machen, Susan R Petrie, and Angela M Ritenour. 1994. The pittsburgh sleep diary. *Journal of sleep research*, 3(2):111–120.
- Charles M Morin, Richard R Bootzin, Daniel J Buysse, Jack D Edinger, Colin A Espie, and Kenneth L Lichstein. 2006. Psychological and behavioral treatment of insomnia: update of the recent evidence (1998–2004). *Sleep*, 29(11):1398–1414.
- Charles M Morin and Colin A Espie. 2007. *Insomnia: A clinical guide to assessment and treatment*. Springer Science & Business Media.
- Jason C Ong, Rachel Manber, Zindel Segal, Yinglin Xia, Shauna Shapiro, and James K Wyatt. 2014. A randomized controlled trial of mindfulness meditation for chronic insomnia. *Sleep*, 37(9):1553–1563.
- Michael L Perlis, Carla Jungquist, Michael T Smith, and Donn Posner. 2006. *Cognitive behavioral treatment of insomnia: A session-by-session guide*, volume 1. Springer Science & Business Media.
- Maria Pontiki, Dimitris Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. **SemEval-2014 task 4: Aspect based sentiment analysis**. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 27–35, Dublin, Ireland. Association for Computational Linguistics.
- Salman Razzaki, Adam Baker, Yura Perov, Katherine Middleton, Janie Baxter, Daniel Mullarkey, Davinder Sangar, Michael Taliercio, Mobasher Butt, Azeem Majeed, et al. 2018. A comparative study of artificial intelligence and human doctors for the purpose of triage and diagnosis. *arXiv preprint arXiv:1806.10698*.
- Steven R Rick, Aaron Paul Goldberg, and Nadir Weibel. 2019. Sleepbot: encouraging sleep hygiene using an intelligent chatbot. In *Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion*, pages 107–108.
- Lee M Ritterband, Frances P Thorndike, Linda A Gonder-Frederick, Joshua C Magee, Elaine T Bailey, Drew K Saylor, and Charles M Morin. 2009. Efficacy of an internet-based behavioral intervention for adults with insomnia. *Archives of general psychiatry*, 66(7):692–698.
- Thomas Roth. 2007. Insomnia: definition, prevalence, etiology, and consequences. *Journal of clinical sleep medicine*, 3(5 suppl):S7–S10.
- Heereen Shim, Stijn Luca, Dietwig Lowet, and Bart Vanrumste. 2020. Data augmentation and semi-supervised learning for deep neural networks-based text classifier. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, pages 1119–1126.
- Arthur J Spielman, Chien-Ming Yang, and Paul B Glovinsky. 2011. Sleep restriction therapy. In *Behavioral treatments for sleep disorders*, pages 9–19. Elsevier.
- Melissa Kaleta Stoller. 1994. Economic effects of insomnia. *Clinical Therapeutics: The International Peer-Reviewed Journal of Drug Therapy*.
- Lars Ström, Richard Pettersson, and Gerhard Andersson. 2004. Internet-based treatment for insomnia: a controlled evaluation. *Journal of consulting and clinical psychology*, 72(1):113.
- Norah Vincent and Samantha Lewycky. 2009. Logging on for better sleep: Rct of the effectiveness of online treatment for insomnia. *Sleep*, 32(6):807–815.
- Aliza Werner-Seidler, Bridianne O’Dea, Fiona Shand, Lara Johnston, Anna Frayne, Andrea S Fogarty, and Helen Christensen. 2017. A smartphone app for adolescents with sleep disturbance: development of the sleep ninja. *JMIR mental health*, 4(3):e28.
- Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. 2019. Huggingface’s transformers: State-of-the-art natural language processing. *ArXiv*, pages arXiv–1910.

A Dataset for pilot study

We collected three data sets for experiment: sleep causes, issues, and impact dataset. Class labels and label descriptions of causes, issues, and impact dataset are summarised in tables 3 to 5, respectively. Each data points annotated with either one or more class labels (max. 3 classes).

B Implementation and training settings

For experiment, PyTorch version (Wolf et al., 2019) of a Bidirectional embedding representations from transformers (BERT) model (Devlin et al., 2019) was used. We initialised the model with pre-trained weights (`bert-base-uncased`) obtained from language modelling with general copora (e.g., Wikicorpus, etc). For classification task, we added a

| Class | Description |
|-------------------|---------------------------|
| exercise | Work out/exercise |
| passivity | Physically not active |
| sleepEnvironment | Bad sleep environment |
| bedTimes | Irregular bed times |
| napping | Nap during the day |
| notTired | Not tired at night |
| breathingIssues | Stopped breathing |
| otherHealthIssues | Other health issues |
| obligationDuties | Too many duties |
| stressMoodAnxiety | Experience stress |
| media | Engage in screen-time |
| caffeine | Consume caffeine |
| pets | Sleep disturbance by pets |
| kids | Sleep disturbance by kids |
| eating | Eat heavy meals |
| drinking | Consume drink |
| alcohol | Consume alcohol drinks |
| nicotine | Smoke |
| noCause | No causes |

Table 3: Class categories for causes dataset

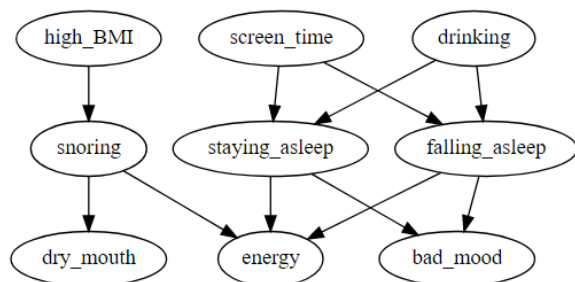


Figure 3: A simple Bayesian Network modelled connectivity between sleep causes, issues, and impact.

final dense layer with sigmoid activation function and used binary cross entropy loss to perform multi-label classification. Details of fine-tuning training are summarised in Table 6.

C Bayesian Network

A Bayesian Network used in a pilot study (Section 4) is illustrated in Figure 3.

| Class | Description |
|-------------------------|--------------------|
| snoringBothersMe | Snoring issue 1 |
| snoringBothersOthers | Snoring issue 2 |
| snoringStoppedBreathing | Stopped breathing |
| staysUpLate | Stay up late |
| troubleFallingAsleep | Lie in bed awake |
| troubleStayingAsleep | Wake up frequently |
| wakeUpTooEarly | Wake up too early |
| problemWakingUp | Trouble waking up |
| sleepsInLater | Sleep in late |
| otherIssue | Other issue |
| goodSleep | No issue |

Table 4: Class categories for issues dataset

| Class | Description |
|----------------------|---------------------------|
| embarrassedBySnoring | Snoring impact |
| dryMouth | Cause dry mouth |
| energy | Feel tired or less energy |
| performance | Affect performance |
| appearance | Look tired |
| stressMoodAnxiety | Bad mood |
| lessPatience | Become less patience |
| socialImpact | Affect social life |
| otherHealthImmunity | Affect health |
| otherImpact | Other impact |
| noImpact | No impact |

Table 5: Class categories for impacts dataset

| Hyperparameter | Assignment |
|----------------------|-------------|
| number of epochs | 4 |
| batch size | 32 |
| learning rate | $5e - 5$ |
| classification layer | feedforward |
| drop out | 0.1 |
| optimizer | AdamW |

Table 6: Hyperparameters for BERT text classifier