

# Toward a Reinforcement-Learning-Based System for Adjusting Medication to Minimize Speech Disfluency

Pavlos Conostas<sup>†</sup>, Vikram Rawal<sup>†</sup>, Matthew Honorio Oliveira<sup>†</sup>, Andreas Conostas<sup>†</sup>, Aditya Khan<sup>†</sup>, Kaison Cheung<sup>†</sup>, Najma Sultani<sup>†</sup>, Carrie Chen<sup>†</sup>, Micol Altomare<sup>†</sup>, Michael Akzam<sup>†</sup>, Jiacheng Chen<sup>†</sup>, Vhea He<sup>†</sup>, Lauren Altomare<sup>†</sup>, Heraa Muqri<sup>†</sup>, Asad Khan<sup>†</sup>, Nimit Amikumar Bhanshali<sup>†</sup>, Youssef Rachad<sup>†</sup> and Michael Guerzhoy<sup>\*</sup>

University of Toronto

## Abstract

We propose a reinforcement learning (RL)-based system that would automatically prescribe a hypothetical patient medication that may help the patient with their mental health-related speech disfluency, and adjust the medication and the dosages in response to zero-cost frequent measurement of the fluency of the patient. We demonstrate the components of the system: a module that detects and evaluates speech disfluency on a large dataset we built, and an RL algorithm that automatically finds good combinations of medications. To support the two modules, we collect data on the effect of psychiatric medications for speech disfluency from the literature, and build a plausible patient simulation system. We demonstrate that the RL system is, under some circumstances, able to converge to a good medication regime. We collect and label a dataset of people with possible speech disfluency and demonstrate our methods using that dataset. Our work is a proof of concept: we show that there is promise in the idea of using automatic data collection to address speech disfluency.

## Keywords

disfluency, ASR, reinforcement learning

## 1. Introduction

Speech disfluency is a common medical issue. It can be caused by, among other factors, conditions such as depression, anxiety, and insomnia (see Section 6). Speech disfluency includes stuttering as well as issues like pauses that are too long, repetitions, “false starts,” and “repairs” of previous utterances [1].

In this paper, we propose a hypothetical Reinforcement Learning-based system for helping physicians ad-

just medications for people with speech disfluency. In the paper, we focus on a proof-of-concept of the system. Of course, regulatory and clinical issues would need to be addressed before actually implementing such a system. However, we show that such a system can in principle be implemented, and automatic disfluency detection can work well enough to support such a system.

On a high level, the system works as follows: the person with a speech disfluency interacts with a system that measures their speech fluency; the system tries to assign different medications in different doses to the person, to minimize their speech disfluency. Using Reinforcement Learning (RL), the system trades off exploring to find the best medication combination for the person and exploiting effective medication combinations that have been found.

We demonstrate two components of the system: a subsystem for detecting how disfluent the person’s speech is, and a subsystem for minimizing the speech disfluency by finding a combination of medications that works using RL.

We train our disfluency detection system to predict the labels we assigned to clips in the dataset we collected.

To demonstrate the feasibility of the RL subsystem, we construct a patient simulation. We measure the precision with which our speech disfluency detection subsystem can measure disfluency, and obtain from the literature the plausible timespans and onset times for the effects of medications. We then run a patient simulation with plausible parameters and demonstrate that our RL subsystem

*Machine Learning for Cognitive and Mental Health Workshop (ML4CMH), AAAI 2024, Vancouver, BC, Canada*

<sup>\*</sup>Corresponding author.

<sup>†</sup>These authors contributed equally.

✉ pconstasm@gmail.com (P. Conostas);  
vikram.rawal@mail.utoronto.ca (V. Rawal);  
matthewhonorio.oliveira@mail.utoronto.ca (M. H. Oliveira);  
andreas.constas@mail.utoronto.ca (A. Conostas);  
aditya.khan@mail.utoronto.ca (A. Khan);  
siukai.cheung@mail.utoronto.ca (K. Cheung);  
najma.sultani@mail.utoronto.ca (N. Sultani);  
carrie.chen71@gmail.com (C. Chen);  
micol.altomare@mail.utoronto.ca (M. Altomare);  
michael.akzam@mail.utoronto.ca (M. Akzam);  
jiachengjason.chen@mail.utoronto.ca (J. Chen);  
vhea.he@mail.utoronto.ca (V. He);  
lauren.altomare@mail.utoronto.ca (L. Altomare);  
heraa.muqri@mail.utoronto.ca (H. Muqri);  
asadk.khan@mail.utoronto.ca (A. Khan);  
nimit.bhanshali@mail.utoronto.ca (N. A. Bhanshali);  
youssef.rachad@mail.utoronto.ca (Y. Rachad);  
guerzhoy@cs.toronto.edu (M. Guerzhoy)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

can find strategies to minimize the speech disfluency of our plausibly-simulated patients.

To evaluate our subsystems, we collect a dataset of public videos of people with possible disfluencies and label the dataset using a scalable strategy that allows us to obtain precise and standardized ratings by having each video be rated by multiple raters.

The rest of the paper is organized as follows: we explain our data collection and labelling process. We then describe our disfluency rating process, and report results on that subsystem. We then describe our patient simulation process and report results of the RL system’s performance on the simulated patients. For the patient simulation to be plausible, we connect the patient simulation to how precisely speech can be rated for fluency by our system and to the plausible effects of medications. (Note that if the measurement of speech fluency is too noisy and/or the medications’ effect is too subtle or the onset is too long, learning would likely not be possible.) Finally, we summarize our literature search results for medications that could plausibly affect speech fluency.

## 2. Data Collection

### 2.1. Methods

The objective of our data collection process is to obtain a series of audio samples from individuals with possible mental health-related speech disfluency, across a period of time. We collected 19 channels from searching on YouTube for mental health-related vlog channels by YouTubers, as well as the D-vlog, a dataset of channels of YouTubers with depression [2].

For each YouTuber represented in the videos, we scraped their channel for other videos which contained significant stretches of unedited spoken audio. Terms used to query for videos from each channel were subsets of the following keywords {“depression”, “story”, “vlog”, “depression vlog”, “anxiety”, “tested”, “figure”, “rambling”, “issues”, “anxiety vlog”, “webcam”}. For each video, only the audio was extracted. In total, we obtained 195 audio clips. There are 9 to 11 audio clips for each channel, with an average of 10 audio clips per channel.

### 2.2. Rating System

We devised a rating system to assess the severity of the disfluency in the video data. The authors acted as raters for the videos. The 19 YouTuber channels in the dataset were examined for disfluencies in them. Raters were tasked with assessing the disfluency severity in each video on a scale of 1 to 7, which was adapted from the Stuttering Severity Instrument-Third Edition (SSI-3) [3]. (But note that “disfluency” is a more general term than “stuttering.”)

The rating process consisted of two stages, each stage lasting approximately one week. In each stage, raters randomly received several channels and were asked to rate their audio samples from the dataset. This was arranged so that each audio sample in the dataset would have up to 3 raters. Each rater received different channels in the different stages. To ensure independent evaluation, raters were advised against sharing their assessments between each other.

Data from the initial stage was not used in our experiments. The round was used for acquainting raters with the variation of disfluency observed in the dataset. At the end of this phase, each rater was privately given summary statistics regarding their ratings in the round, including the mean and standard deviation of their ratings across audio samples, as well as a spreadsheet containing a measure of their bias for each audio sample (where bias is the distance of their rating from the mean rating across all raters for that audio sample). This process was aimed at allowing raters to recognize possible inconsistencies and biases in their “internal model” of disfluency. The ratings from the second stage were the finalized ratings that would be used for fine-tuning our disfluency-detection system. The ratings were standardized, as described below.

## 3. Rater Performance Analysis

In this Subsection, we analyze the rater data, and show that raters are somewhat consistent in their ratings of the same clips. This indicates that we can use the standardized ratings (see below) as targets when estimating the fluency of speakers in audio clips.

### 3.1. Data Model

To assess the performance of the raters, we conducted a regression analysis. The model we utilized was

$$r_{ij} = \alpha_i + \beta_j + \varepsilon_{ij}$$

where  $r_{ij}$  is the rating given to audio clip  $j$  by rater  $i$ ,  $\alpha_i$  is the rater bias,  $\beta_j$  is the true average disfluency of the channel, and  $\varepsilon_{ij}$  is the random error (see [4] for a similar model). This model is estimated using least-squares regression.

Using this model, the performance of the raters was assessed by randomly splitting the dataset into a 70% training set, and a 30% validation set.

### 3.2. Analysis

Below, we perform an exploratory analysis of the non-standardized ratings.

We compute the Root-Mean-Square Error (RMSE) on the training set and the validation set when predicting the disfluency scores using the data model. The RMSE on the training set was 0.8/6.0 (on a scale of 0 to 6 rather than 1 to 7 as in the input) and the RMSE on the validation set was 0.9/6.0. The validation RMSE we would obtain if the data model predicted the average rating every time would be 1.4/6.0. The  $R^2$  value of the model on the training set was 0.44, indicating that the rater coefficients and the clip coefficients have explanatory power.

For each clip on the validation set, we compute the standard deviation of the ratings assigned by different raters to the same clip. The median standard deviation is 0.6/6.0. This suggests that the median disagreement between raters was just over half a rating point on a given clip. The standard deviations are given on a scale of 6.0 since the scores range from 1 to 7.

### 3.3. Standardized Ratings

Different raters use different standards for fluency. We therefore obtained *standardized* ratings. We accomplish this by subtracting the rater bias  $\alpha_i$  (see Section 3.1) for rater  $i$  for each rating  $r_{ij}$  by rater  $i$ . Then, when we compute the average standardized rating for every clip, we average ratings that are actually on the same scale.

## 4. Disfluency Pipeline

In this section, we describe our subsystem for assessing the disfluency of a person in the input clip. We use Whisper<sup>1</sup> [5] to transcribe the audio. We then use an Auto-Correlational Neural Network-based tagger [6] to tag the Whisper transcript. Finally, we fine-tune GPT-2 [7] on the tagged transcript as input in order to predict the disfluency scores we assigned.

### 4.1. Transcribing Audio with Whisper

The YouTube videos are transcribed using the Automated Speech Recognition (ASR) model Whisper. Tokens such as “uh”, “um”, etc. were included in the transcript.

### 4.2. Disfluency Tagging

The parsed text transcripts were subsequently fed into a Disfluency Tagging Auto-Correlational Neural Network (DT-ACNN) [6] – a system designed to categorize each word within the text transcript as either “fluent” or “disfluent”. In [6], the Switchboard corpus of conversational speech [8] dataset was used. For the task of predicting a per-word “fluent” or “disfluent” label, the authors report a recall of 90.0%, a precision of 82.8%, and an F1 score

<sup>1</sup><https://github.com/openai/whisper>

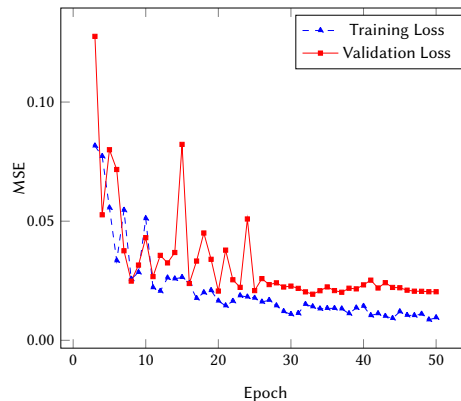


Figure 1: Learning curves: disfluency prediction regression task

of 86.2% on the dataset. The reported results indicate the effectiveness of the DT-ACNN model in disfluency detection.

### 4.3. Fine-tuning GPT-2

We fine-tune GPT-2 to predict the average standardized disfluency score by the rates who rated the clip from the tagged Whisper transcripts, as well as from the words-per-minute (WPM) measure.

For the regression task, we train with embedding size 768, using the Mean Squared Error (MSE) loss, and the AdamW optimizer with parameters  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\epsilon = 10^{-9}$ . The token limit of GPT2 is 1024. Inputs that exceed this limit were truncated. The following hyper-parameters were used during training: a learning rate of  $4.5 \times 10^{-4}$ , a batch size of 4 (dictated by computational limitations), with weight decay parameter 0.01, for 50 epochs.

P-tuning [9] was used. In this approach, a soft prompt with a set of 100 tokens is introduced at the beginning of the input. These tokens aid in guiding the model during classification. The model uses a Prompt Encoder to optimize the prompt, with an encoding layer comprised of 128 units. The model performance was evaluated by randomly splitting the dataset into an 80% training set, and a 20% validation set.

### 4.4. Results

The learning curves for the disfluency prediction task are in Fig. 1. We observe that our system is currently able to predict the validation rating to within about 0.15/6 of the actual rating on average (the standard error is obtained by taking the square root of the MSE) for YouTubers not in the training set.

## 5. Patient Simulation and Reinforcement Learning

### 5.1. Overview

In this Section, we explore the plausibility of using RL together with signals from our speech disfluency detector to find an effective medication regimen for people with speech disfluency.

We first describe how we simulate people with speech disfluency in a plausible way. We then demonstrate that our RL algorithm could find an effective medication regimen in a plausible scenario.

### 5.2. Prior Work: RL for Medication Adjustment

Reinforcement Learning (RL) for medication adjustment has been proposed in several contexts. Oh et al. evaluate an offline RL algorithm learned on South Korea's national Health insurance system to prescribe diabetes medication [10]. Javad et al. similarly propose an offline RL algorithm [11]. They measure the performance of the system based on the concordance to the prescription actually made, as well as by analyzing outcomes where the system's recommendation and the actual recommendation in the data disagreed. Sun et al. explored an approach for Type 2 Diabetes treatment [12]. They merged a knowledge-driven model, informed by clinical guidelines, with a data-driven deep reinforcement learning model. The knowledge-driven model uses data from the Singapore Health Services Diabetes Registry which contains over 189,520 patients and their Type-2 Diabetes medication prescription to narrow down a list of viable medications to which the data-driven model applies a Deep Q Network (DQN) which also learns from the historical patient data and is used to rank the candidate medications selected by the knowledge-driven model based on expected long-term rewards.

Nemati et al. developed a clinician-in-the-loop framework for heparin dosing, leveraging data from the MIMIC-II intensive care unit database [13]. This study engaged an interactive agent in simulated dosing trials, learning from the outcomes to refine decision-making processes. Similarly, Anzabi Zadeh et al. utilized deep reinforcement learning in the context of warfarin dosing for patients with blood clotting issues with an emphasis on individualized dosing due to warfarin's narrow therapeutic range [14]. In this method, they frame the problem as a Markov decision process (MDP) and employ an agent within a Pharmacokinetic/Pharmacodynamic (PK/PD) model to simulate dose-responses of virtual patients in which the agent learns the best dose-duration pair through experience replay.

### 5.3. Person with Disfluency Simulation

We model a medication administration environment, aiming to determine the most effective medication regime for people experiencing depression, anxiety, insomnia, and resulting speech fluency issues. The person's health state evolves based on Hidden Markov Models (HMMs). Each health issue (depression, anxiety, insomnia) has its unique HMM, governing how the patient's state progresses. The patient's observed speech fluency is also influenced by these health states.

The Medication object represents different types of medications, each with varying effects on the aforementioned health issues. These effects include beneficial impacts on the conditions and potential side effects. Medications have properties like dosage, half\_life, and time\_to\_effect, which dictate how they function over time.

We simulate people with disfluency by evolving the HMM state. The patient model has the following attributes:

- *Depression, Anxiety, Insomnia Scores*: These attributes represent the initial underlying conditions of the patient. Represented as an integer between 1 and 5. A higher number denotes a more severe state.
- *Depression, Anxiety, Insomnia Hidden Markov Models*: Models that represent the behaviour of how the severity of the patient's depression, anxiety, and insomnia change over time based on their initial states and also through interaction with medicine. These directly impact observed speech fluency context.
- *Speech Fluency Score*: Indicates the patient's natural ability to speak fluently, modelled as a continuous value between 0 and 1.
- *Medication Accumulation*: A list that keeps track of all medications that are currently in the patient's system.

Alongside this, we also model an individual medication with the following attributes:

- *Name*: The name of the medication.
- *Depression, Anxiety, Insomnia Effects*: Captures the medication's average effect and variability on each condition given the standard dosage.
- *Dosage*: The amount of medication administered relative to the standard dose. (e.g. Dosage = 1.5 means 1.5x the standard dose). This attribute scales the effects of the medication on the patient.
- *Time to Effect*: The amount of days it takes for the medication to start showing effects.
- *Half-Life*: The amount of days it takes for the medication dosage in the patient's system to reduce by half.

## 5.4. Hidden Markov Model (HMM)

A Hidden Markov Model (HMM) is a statistical model that represents sequences of observable data as well as hidden states. The sequences of observable data are generated based on hidden states which cannot be directly observed. Here, the “observable data” is the patient’s speech fluency score while the “hidden states” are the underlying depression, anxiety, and insomnia conditions that affect the severity of the disfluency. We base this model off of the fact that a patient’s underlying level of depression [15], insomnia [16], and anxiety [17] have an impact on their speech fluency.

Each health condition — depression, anxiety, and insomnia — has its associated HMM. The key components of these HMMs are:

- The initial probability distribution over the initial state of the condition. Initialized as a uniform distribution, indicating that any severity level is equally likely at the start.
- The transition matrix, which defines the probability of transitioning from one state (severity level) to another in consecutive time steps. For instance, if a patient is currently at a severity level of 3 for depression, the transition matrix, the transition matrix will dictate the probability of them improving to level 2, worsening to level 4, or remaining at level 3 in the next step.
- We use a Gaussian Hidden Markov Model as the observable context is assumed to be generated from a Gaussian distribution. The means and covariances define these distributions for each hidden state.
- Each state has a mean context emission, set to be the same in the depression, anxiety, and insomnia states.

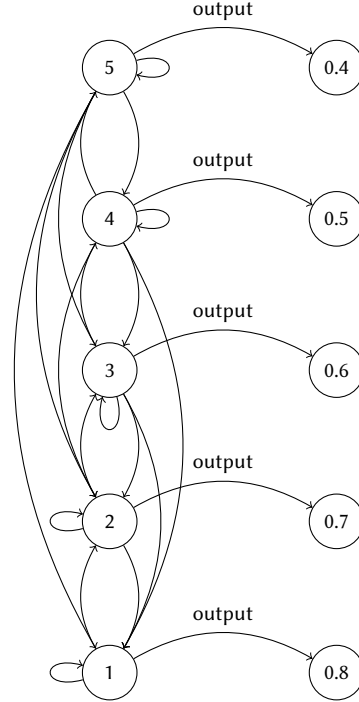
See Fig. 2 for a diagram.

## 5.5. RL Environment

At each step, an agent can choose to administer a specific medication from the available list. The environment then evolves based on the medication’s effects and the underlying psychiatric state of the patient model by the use of a transition matrix that map the current psychiatric state to a new state based on a probability distribution that models the dynamic and evolving nature of underlying psychiatric states [18]. The agent receives a reward based on the patient’s measured fluency.

We use the LinUCB [19] algorithm to learn the optimal medication strategy. The goal is to maximize the patient’s speech fluency.

The effects of the medication on the patient model is implemented by applying the effects of the medication on each condition on each condition’s transition matrix.



**Figure 2:** A diagram of a Hidden Markov Model representing context emissions generated from states.

## 5.6. Medication Selection Algorithm

We use the increase in speech fluency as our reward. Speech fluency is modelled as a linear function:

$$S = 0.1 \cdot D + 0.2 \cdot A + 0.3 \cdot I + 0.4 \cdot R$$

where  $S$  is current speech fluency,  $D$ ,  $A$ ,  $I$  are the patient’s current depression, anxiety, and insomnia scores respectively, labelled on a 5 point scale, where 1 represents no symptoms and 5 represents the most severe symptoms.  $R$  is the patient’s baseline fluency.

The implementation of the LinUCB algorithm observes the current state of the patient, then for each medication that is part of the environment, it estimates the reward using a linear approximation. An upper confidence bound is calculated for the estimated reward in which the medication with the highest upper confidence bound is chosen according to the equation:

$$a_t = \arg \max_{a \in A} \left( x_t(a)^T \theta_a + \alpha \sqrt{x_t(a)^T A_a^{-1} x_t(a)} \right)$$

where  $x_t(a)$  is the feature vector for action  $a$  at time  $t$ ,  $\theta_a$  is the parameter vector for action  $a$  which we want to estimate, and  $\alpha$  is the hyperparameter controlling the

exploration/exploitation trade-off [20]. In our implementation of the LinUCB algorithm, we chose the value of  $\alpha = 10.0$ .

## 5.7. Results

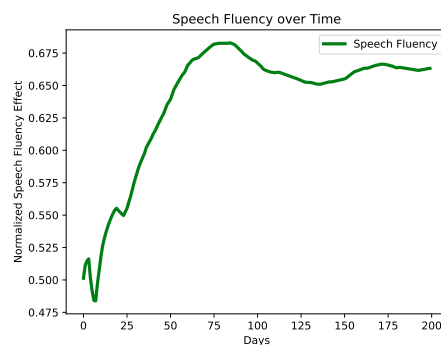
In our simulations, we run the RL algorithm and keep track of disfluency over time. We inject noise into the algorithm’s simulated measurements of disfluency to simulate the fact that our disfluency detection system does not measure disfluency perfectly. In the experiments reported here, we inject a minimal amount of noise, corresponding to the high precision with which we can measure disfluency.

We define the success of a trial as an improvement of over  $0.5\sigma$  in speech fluency. We define failure as a deterioration of over  $0.5\sigma$  in speech fluency. Here,  $\sigma \approx 0.1$  is the standard deviation of fluency in the dataset.

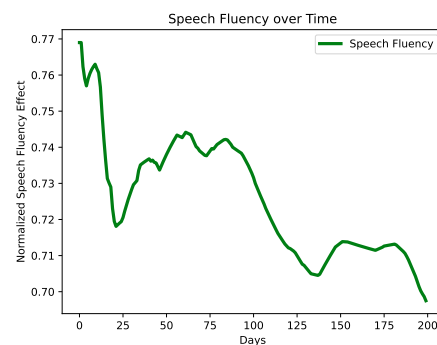
Figures 3 and 4 show examples of a successful and an unsuccessful run, respectively. Across 500 patient simulation runs, we found a high potential for reinforcement learning to correctly apply medication effects to reduce speech disfluency, with 52% of runs showing success and 9% of runs demonstrating failure.

The average fluency across these runs was 0.66/1.00 with a standard deviation of 0.1. The success rate of the simulations was 52%, with a failure rate of approximately 16%. Success and failure are defined as runs terminating greater than  $0.5\sigma$  and lower than  $-0.5\sigma$  from the initial fluency level, respectively.

These results support the possibility of the use of reinforcement learning to improve speech fluency under the studied conditions. Our preliminary results indicate that if speech disfluency can be measured to within 10% of the true score and the medications have plausible properties (similar to the ones seen in Section 6), then reinforcement learning is a possible method to dose medications that pessimize speech disfluency. However, the variability in outcomes and the presence of failed simulations that prompted theoretical patient deterioration indicate that further research is needed in improving the model’s accuracy and understanding factors contributing to failures, that will be important for applying these findings in a clinical setting.



**Figure 3**  
Convergence to higher speech fluency.



**Figure 4**  
Lack of convergence to higher speech fluency.

## 6. Medication Literature Review

A systematic literature review was conducted to determine common medications used to treat major depressive disorder (“depression”) and other mental illnesses that affect speech fluency. Table 1 indicates 23 medications whose onset time and response rates were used to inform the reinforcement learning simulation.

## 7. Ethical Considerations

In this paper, we outline and evaluate a proposal to adjust medications automatically in order to improve the speech fluency of a simulated patient. Although administering medications automatically is sometimes done (e.g., with Insulin pumps), this can only be done after thorough clinical trials and with informed consent from the patient. Patients must be thoroughly informed about the nature of the automated system, its potential risks and benefits, and their rights in the decision-making progress. This extends beyond initial consent, and includes ongo-

ing consent as the system adapts medication regimens. We believe that if safe dosage ranges can be determined, automatically adjusting medications can in principle be done. Potential upsides include better-optimized medication regimens that are found faster (or found at all). The system’s reliance on sensitive patient data raises significant privacy and security concerns. Finally, the long-term implications of integrating such a system into healthcare practices must be considered. This includes ongoing monitoring of the system’s impact on patient health outcomes, adapting to new medical insights, and updating the system accordingly.

## 8. Limitations

This paper studies the feasibility of using automated measures of disfluency together with a reinforcement learning system to adjust medications. Our model is informed by the literature, but it is still a “toy” model. However, our results indicate that it is possible to measure disfluency to a certain precision (for some definition of “disfluency”), and that this precision can be sufficient to serve as input to an RL system.

## 9. Conclusions and Future Work

We propose and evaluate a new idea: people with speech disfluency can have their disfluency measured automatically, and a reinforcement learning algorithm can find an optimal medication regime for them through exploration/exploitation. We have shown that the components of this system are possible to build.

In our preliminary experiments, we have shown that with an accurate enough disfluency detection system and with medications with plausible properties, a good medication regime can be found by a reinforcement learning system.

Future work includes improving our subsystems for performance, ensuring that our simulations preserve patient safety (e.g. by only trying safe dosages), and running more extensive simulations to show under what conditions RL for optimizing medications is feasible.

## References

- [1] E. Shriberg, R. Bates, A. Stolcke, A prosody only decision-tree model for disfluency detection, in: *Fifth European Conference on Speech Communication and Technology*, 1997.
- [2] J. Yoon, C. Kang, S. Kim, J. Han, D-vlog: Multimodal vlog dataset for depression detection, *Proceedings of the AAAI Conference on Artificial Intelligence* 36 (2022) 12226–12234.
- [3] R. Glyndon, *Stuttering severity instrument for children and adults*. 3rd ed., Pro-Ed, Austin, Texas, 1994.
- [4] M. R. Raymond, C. Viswesvaran, Act research report series: Least-squares models to correct for rater effects in performance assessment, *PsycEXTRA Dataset* (1991). doi:10.1037/e426952008-001.
- [5] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, I. Sutskever, Robust speech recognition via large-scale weak supervision, in: *International Conference on Machine Learning*, PMLR, 2023, pp. 28492–28518.
- [6] P. J. Lou, P. Anderson, M. Johnson, Disfluency detection using auto-correlational neural networks, 2020. arXiv:1808.09092.
- [7] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever, et al., Language models are unsupervised multitask learners, *OpenAI blog* 1 (2019) 9.
- [8] J. J. Godfrey, E. C. Holliman, J. McDaniel, Switchboard: Telephone speech corpus for research and development, in: *Acoustics, speech, and signal processing, IEEE international conference on*, volume 1, IEEE Computer Society, 1992, pp. 517–520.
- [9] X. Liu, Y. Zheng, Z. Du, M. Ding, Y. Qian, Z. Yang, J. Tang, Gpt understands, too, *AI Open* (2023). doi:10.1016/j.aiopen.2023.08.012.
- [10] S. H. Oh, J. Park, S. J. Lee, S. Kang, J. Mo, Reinforcement learning-based expanded personalized diabetes treatment recommendation using south korean electronic health records, *Expert Systems with Applications* 206 (2022) 117932.
- [11] M. O. M. Javad, S. O. Agboola, K. Jethwani, A. Zeid, S. Kamarthi, et al., A reinforcement learning-based method for management of type 1 diabetes: exploratory study, *JMIR diabetes* 4 (2019) e12905.
- [12] X. Sun, Y. M. Bee, S. Lam, Z. Liu, W. Zhao, S. Y. Chia, H. Kadir, J. Wu, B. Ang, N. Liu, Z. Lei, Z. Xu, T. Zhao, G. Hu, G. Xie, Effective treatment recommendations for type 2 diabetes management using reinforcement learning: Treatment recommendation model development and validation, *Journal of Medical Internet Research* 23 (2021) e27858. doi:10.2196/27858.
- [13] S. Nemati, M. M. Ghassemi, G. D. Clifford, Optimal medication dosing from suboptimal clinical examples: A deep reinforcement learning approach, 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (2016) 2978–2981. URL: <https://api.semanticscholar.org/CorpusID:10907669>.
- [14] S. Anzabi Zadeh, W. N. Street, B. W. Thomas, Optimizing warfarin dosing using deep reinforcement learning, *Journal of Biomedical Informatics* 137 (2023) 104267. URL: <http://dx.doi.org/10.1016/j.jbi.2022.104267>. doi:10.1016/j.jbi.2022.104267.

- [15] P. Fossati, L. Bastard Guillaume, A.-M. Ergis, J.-F. Allilaire, Qualitative analysis of verbal fluency in depression, *Psychiatry Research* 117 (2003) 17–24. URL: <https://www.sciencedirect.com/science/article/pii/S0165178102003001>. doi:[https://doi.org/10.1016/S0165-1781\(02\)00300-1](https://doi.org/10.1016/S0165-1781(02)00300-1).
- [16] M. M. Jacobs, S. Merlo, P. M. Briley, Sleep duration, insomnia, and stuttering: The relationship in adolescents and young adults, *Journal of Communication Disorders* 91 (2021) 106106. URL: <https://www.sciencedirect.com/science/article/pii/S0021992421000290>. doi:<https://doi.org/10.1016/j.jcomdis.2021.106106>.
- [17] Z. Wang, M. Tang, M. Larrazabal, E. Toner, M. Rucker, C. Wu, B. Teachman, M. Boukhechba, L. Barnes, Personalized state anxiety detection: An empirical study with linguistic biomarkers and a machine learning pipeline (2023). doi:<https://doi.org/10.48550/arXiv.2304.09928>.
- [18] C. Gauld, D. Depannemaecker, Dynamical systems in computational psychiatry: A toy-model to apprehend the dynamics of psychiatric symptoms, *Frontiers in Psychology* 14 (2023). doi:10.3389/fpsyg.2023.1099257.
- [19] E. Nelson, D. Bhattacharjya, T. Gao, M. Liu, D. Bounieffouf, P. Poupert, Linearizing contextual bandits with latent state dynamics, in: *Uncertainty in Artificial Intelligence*, PMLR, 2022, pp. 1477–1487.
- [20] L. Li, W. Chu, J. Langford, R. E. Schapire, A contextual-bandit approach to personalized news article recommendation, in: *Proceedings of the 19th international conference on World wide web*, 2010, pp. 661–670.
- [21] M. Wilson, J. Tripp, *Clomipramine* (2019).
- [22] H. Abdul-Baki, I. I. El Hajj, L. ElZahabi, C. Azar, E. Aoun, A. Skoury, H. Chaar, A. I. Sharara, A randomized controlled trial of imipramine in patients with irritable bowel syndrome, *World journal of gastroenterology: WJG* 15 (2009) 3636.
- [23] C. A. Townsend, Selegiline transdermal patch (emsam) for major depressive disorder, *American Family Physician* 77 (2008) 505.
- [24] J. J. Moore, A. Saadabadi, Selegiline, in: *StatPearls [Internet]*, StatPearls Publishing, 2022.
- [25] T. Tenjin, S. Miyamoto, Y. Ninomiya, R. Kitajima, S. Ogino, N. Miyake, N. Yamaguchi, Profile of bionanserin for the treatment of schizophrenia, *Neuropsychiatric disease and treatment* (2013) 587–594.
- [26] L. Dean, Venlafaxine therapy and cyp2d6 genotype (2020).
- [27] R. D. Gibbons, K. Hur, C. H. Brown, J. M. Davis, J. J. Mann, Benefits from antidepressants: synthesis of 6-week patient-level outcomes from double-blind placebo-controlled randomized trials of fluoxetine and venlafaxine, *Archives of general psychiatry* 69 (2012) 572–579.
- [28] S. L. Cincotta, J. S. Rodefer, Emerging role of sertindole in the management of schizophrenia, *Neuropsychiatric disease and treatment* (2010) 429–441.
- [29] J. S. Maan, T. Duong, A. Saadabadi, Carbamazepine (2018).
- [30] Z. Tolou-Ghamari, M. Zare, J. M. Habibabadi, M. R. Najafi, A quick review of carbamazepine pharmacokinetics in epilepsy from 1953 to 2012, *Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences* 18 (2013) S81.
- [31] A. C. Pande, J. G. Crockatt, D. E. Feltner, C. A. Janney, W. T. Smith, R. Weisler, P. D. Londeborg, R. J. Bielski, D. L. Zimbroff, J. R. Davidson, et al., Pregabalin in generalized anxiety disorder: a placebo-controlled trial, *American Journal of Psychiatry* 160 (2003) 533–540.
- [32] J. R. Strawn, L. Geraciotti, N. Rajdev, K. Clemenza, A. Levine, Pharmacotherapy for generalized anxiety disorder in adult and pediatric patients: an evidence-based treatment review, *Expert opinion on pharmacotherapy* 19 (2018) 1057–1070.
- [33] K. Chokhawala, S. Lee, A. Saadabadi, Lithium. national library of medicine, National Center for Biotechnology Information. PubMed Central (2022).
- [34] T. Hui, A. Kandola, L. Shen, G. Lewis, D. Osborn, J. Geddes, J. Hayes, A systematic review and meta-analysis of clinical predictors of lithium response in bipolar disorder, *Acta Psychiatrica Scandinavica* 140 (2019) 94–115.
- [35] D. F. Ionescu, R. C. Shelton, L. Baer, K. H. Meade, M. B. Swee, M. Fava, G. I. Papakostas, Ziprasidone augmentation for anxious depression, *International clinical psychopharmacology* 31 (2016) 341.
- [36] T. Zhao, T.-W. Park, J.-C. Yang, G.-B. Huang, M.-G. Kim, K.-H. Lee, Y.-C. Chung, Efficacy and safety of ziprasidone in the treatment of first-episode psychosis: an 8-week, open-label, multicenter trial, *International clinical psychopharmacology* 27 (2012) 184–190.
- [37] F.-G. Pajonk, Risperidone in acute and long-term therapy of schizophrenia—a clinical profile, *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 28 (2004) 15–23.
- [38] J. Peuskens, Risperidone in the treatment of patients with chronic schizophrenia: a multi-national, multi-centre, double-blind, parallel-group study versus haloperidol, *The British Journal of Psychiatry* 166 (1995) 712–726.
- [39] M. J. Allen, S. Sabir, S. Sharma, Gaba receptor (2018).
- [40] M. Panebianco, S. Al-Bachari, J. L. Hutton, A. G. Marson, Gabapentin add-on treatment for drug-resistant focal epilepsy, *Cochrane Database of Systematic Reviews* (2021).



- [41] F. Lavergne, I. Berlin, A. Gamma, H. Stassen, J. Angst, Onset of improvement and response to mirtazapine in depression: a multicenter naturalistic study of 4771 patients, *Neuropsychiatric disease and treatment* 1 (2005) 59–68.
- [42] H. R. Song, W.-M. Bahk, Y. S. Woo, J.-H. Jeong, Y.-J. Kwon, J. S. Seo, W. Kim, M.-D. Kim, Y.-C. Shin, S.-Y. Lee, et al., Efficacy and tolerability of generic mirtazapine (mirtax) for major depressive disorder: multicenter, open-label, uncontrolled, prospective study, *Clinical Psychopharmacology and Neuroscience* 13 (2015) 144.
- [43] K. A. Fariba, A. Saadabadi, *Topiramate* (2020).
- [44] Y.-T. Liu, G.-T. Chen, Y.-C. Huang, J.-T. Ho, C.-C. Lee, C.-C. Tsai, C.-N. Chang, Effectiveness of dose-escalated topiramate monotherapy and add-on therapy in neurosurgery-related epilepsy: A prospective study, *Medicine* 99 (2020).
- [45] G. Lewis, L. Duffy, A. Ades, R. Amos, R. Araya, S. Brabyn, K. S. Button, R. Churchill, C. Derrick, C. Dowrick, et al., The clinical effectiveness of sertraline in primary care and the role of depression severity and duration (panda): a pragmatic, double-blind, placebo-controlled randomised trial, *The Lancet Psychiatry* 6 (2019) 903–914.
- [46] M. F. Flament, R. Lane, R. Zhu, Z. Ying, Predictors of an acute antidepressant response to fluoxetine and sertraline, *International clinical psychopharmacology* 14 (1999) 259–276.
- [47] H. K. Singh, A. Saadabadi, *Sertraline* (2019).
- [48] M. H. Trivedi, A. J. Rush, S. R. Wisniewski, A. A. Nierenberg, D. Warden, L. Ritz, G. Norquist, R. H. Howland, B. Lebowitz, P. J. McGrath, et al., Evaluation of outcomes with citalopram for depression using measurement-based care in star\* d: implications for clinical practice, *American journal of Psychiatry* 163 (2006) 28–40.
- [49] Z. Jia, J. Yu, C. Zhao, H. Ren, F. Luo, Outcomes and predictors of response of duloxetine for the treatment of persistent idiopathic dentoalveolar pain: A retrospective multicenter observational study, *Journal of Pain Research* (2022) 3031–3041.
- [50] N. Parikh, M. Yilanli, A. Saadabadi, *Tranlycypromine* (2017).
- [51] W. T. Heijnen, J. De Fruyt, A. I. Wierdsma, P. Sien-aert, T. K. Birkenhäger, Efficacy of tranlycypromine in bipolar depression: a systematic review, *Journal of clinical psychopharmacology* 35 (2015) 700–705.
- [52] J. Waugh, K. L. Goa, Escitalopram: a review of its use in the management of major depressive and anxiety disorders, *CNS drugs* 17 (2003) 343–362.
- [53] K. Jiang, L. Li, X. Wang, M. Fang, J. Shi, Q. Cao, J. He, J. Wang, W. Tan, C. Hu, Efficacy and tolerability of escitalopram in treatment of major depressive disorder with anxiety symptoms: a 24-week, open-label, prospective study in chinese population, *Neuropsychiatric Disease and Treatment* (2017) 515–526.
- [54] J. Cookson, P. E. Keck Jr, T. A. Ketter, W. Macfadden, Number needed to treat and time to response/remission for quetiapine monotherapy efficacy in acute bipolar depression: evidence from a large, randomized, placebo-controlled study, *International clinical psychopharmacology* 22 (2007) 93–100.
- [55] J.-S. Lee, J.-H. Ahn, J.-I. Lee, J.-H. Kim, I. Jung, C.-U. Lee, J.-Y. Lee, S.-I. Lee, C.-Y. Kim, Dose pattern and effectiveness of paliperidone extended-release tablets in patients with schizophrenia, *Clinical Neuropharmacology* 34 (2011) 186–190.
- [56] A. Kumar, S. Balan, Fluoxetine for persistent developmental stuttering, *Clinical neuropharmacology* 30 (2007) 58–59.

Drug Name	Medication Onset Time	Variance in Patient Response	Treated Condition
Clomipramine	6-12 weeks [21]	Low remission rate (20%); [21]	Obsessive compulsive disorder [21]
Imipramine	4-5 weeks [22]	80.6% patients after 12-week treatment, compared to 48.0% in the placebo [22]	Depression, anxiety [22]
Selegiline	1-2 weeks [23]	33-40% [23]	Parkinson's disease, depression [24]
Blonanserin	4-6 weeks [25]	52-87% [25]	Schizophrenia, mania [25]
Venlafaxine	4-6 weeks [26]	67.2% [27]	Depression, anxiety, panic disorders [26]
Sertindole	4-6 weeks [28]	85% [28]	Schizophrenia [28]
Carbamazepine	First few days [29]	75-85% [30]	Bipolar disorder, seizures [29]
Pregabalin	3 days [31]	40% [32]	Anxiety [32]
Lithium	1-3 weeks [33]	33% [34]	Bipolar disorder [33]
Ziprasidone	1-2 weeks [35]	51.85% [36]	Bipolar disorder and associated depression [35]
Risperidone	4 weeks [37]	63.4% (4 mg), 65.8% (8 mg) [38]	Treats irritability associated with autism and schizophrenia [38]
Gabapentin	1-4 weeks [39]	17.2% - 37.6% (600 mg-1800mg dosage) [40]	Epilepsy [39]
Mirtazapine	1 week [41]	67.1% [42]	Depression, anxiety, obsessive compulsive disorder [41]
Topiramate	2-4 weeks (epilepsy), 3 months (migraines) [43]	88% [44]	Epilepsy, migraines [43]
Sertraline	Within 6 weeks [45]	59% [46]	Depression, PTSD, OCD, panic disorder, social anxiety disorder [47]
Citalopram	1-2 weeks to start working, 4-6 weeks for full benefit [48]	47% [48]	Depression, social anxiety disorder, PTSD [48]
Duloxetine	2-4 weeks [49]	77% [49]	Depression, anxiety [49]
Tranlycypromine	1-2 weeks to start working, 6-8 weeks for full improvement [50]	60%-80.7% [51]	Major depressive episodes [50]
Escitalopram	1-2 weeks to start working, 6-8 weeks for full improvement [52]	30.2% after 2 weeks of treatment and 75.8% after 8 weeks of treatment [53]	Depression, generalized anxiety disorder (GAD), obsessive compulsive disorder (OCD) and panic attacks [52]
Quetiapine	1-2 weeks to start working, 2-3 months for full improvement [54]	58.2% for quetiapine 600 mg/day and 57.6% for quetiapine 300 mg/day after 8 weeks of treatment [54]	Schizophrenia, manic, psychotic and depressive episodes [54]
Paliperidone	2-8 weeks for full improvement [55]	33.6% after 2 weeks [55]	Psychotic disorders including schizophrenia [55]
Fluoxetine	4-5 weeks [56]	41% [56]	OCD, certain eating disorders, panic attacks [56]

**Table 1**

Summary of medication and variation in patient response for depression and related mental health conditions.