Modeling Mental Health

Wendy Matta, Abhas Wanchu, Michael Garfagnoli, Aimee Williams, Nura Hossainzadeh

Introduction

In this project, we use machine learning models to gauge the mental health of the authors of short statements that were posted in online forums.

Table of contents

01

Motivation

Our emotions are deeply complex, deeply intimate, and central to our humanity. But can they be understood by artificial intelligence?

Motivation: Why is this interesting?

- **From a medical standpoint:** How do words written in online forums give us insight into mental health conditions? Can models help us to predict and prevent more severe future conditions?
- **From a machine learning standpoint:** Are models capable of detecting semantic patterns associated with mental health conditions?
- **From a philosophical standpoint:** Can mathematical algorithms and online written expression help us to understand emotion? Or do we miss something?
- **From a social justice standpoint:** Can machine learning models (or the chatbots powered by them) serve as more unbiased assessors of mental health, unhindered by human prejudice?

What has been done before?

Machine learning has been used...

To predict mental health crises based on electronic health records

Garriga, Roger, et al. "Machine learning model to predict mental health crises from electronic health records." *Nature medicine* 28.6 (2022): 1240-1248.

To diagnose and understand current mental health conditions

Iyortsuun, Ngumimi Karen, et al. "A review of machine learning and deep learning approaches on mental health diagnosis." *Healthcare*. Vol. 11. No. 3. MDPI, 2023.

To develop chatbots that could help with intervention

Abd-Alrazaq, Alaa A., et al. "An overview of the features of chatbots in mental health: A scoping review." *International journal of medical informatics* 132 (2019): 103978.

Our approach to the question

Develop two models: a binary one to predict healthy/not healthy, and a multiclass one to predict mental health status (one of 7 classes)

Baseline models: simple majority class prediction and logistic regression for multiclass prediction

Developed models: Bag of words model, embeddings model, transformer model, model for multi-hot encoded data. Use Keras tuner to determine best parameters.

02

Data

Data Sources

Kaggle

Combination of 9 different datasets

Online Discussion Platforms

Facebook, Reddit, Twitter (X), Social Media

Conversations

Interviews, transcripts of interactions

Mental Health Categorization

Depression, Suicidal, Anxiety, Stress, Bipolar, Personality Disorder, None (Normal)

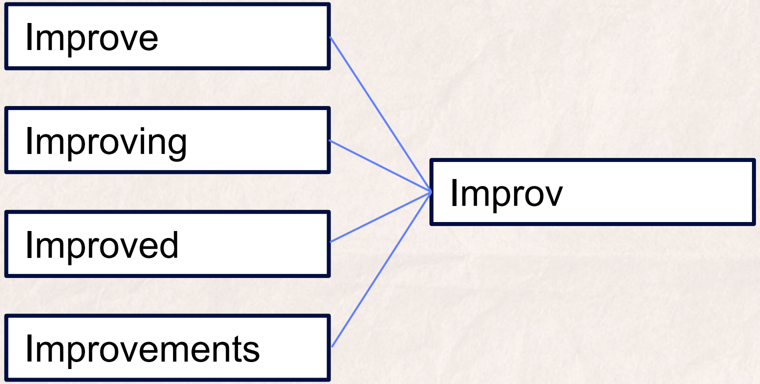
Data Pre-Processing

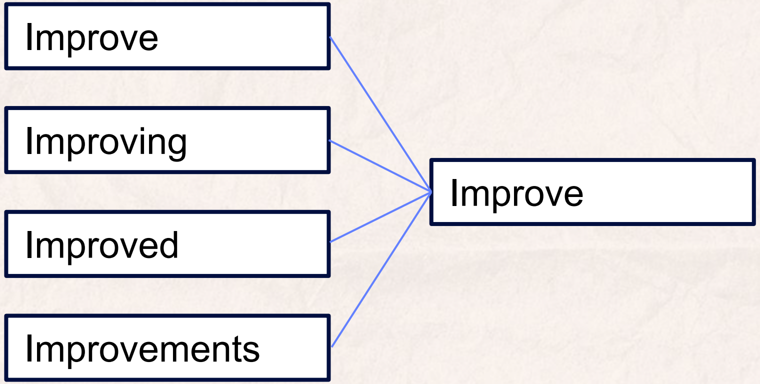
1. Lowercase and Remove Nulls
2. Remove and Expand Contractions
3. Remove Duplicate Text
4. Emoji and Emoticon Translation
5. Remove Stop Words
6. Lemmatization Process

Emoji & Emoticon Translation

Emoji/Emoticon	Text Replacement
:)	smiling_face
:-)	smiling_face
😁	smiling_face
:(	frowning_face
😞	frowning_face

Stemming vs Lemmatization





We chose lemmatization over stemming because the output of stemming is often a root form that may not be a real word (or a word that may have differing context), while lemmatization yields actual words (lemmas).

EDA

Class (Im)Balance

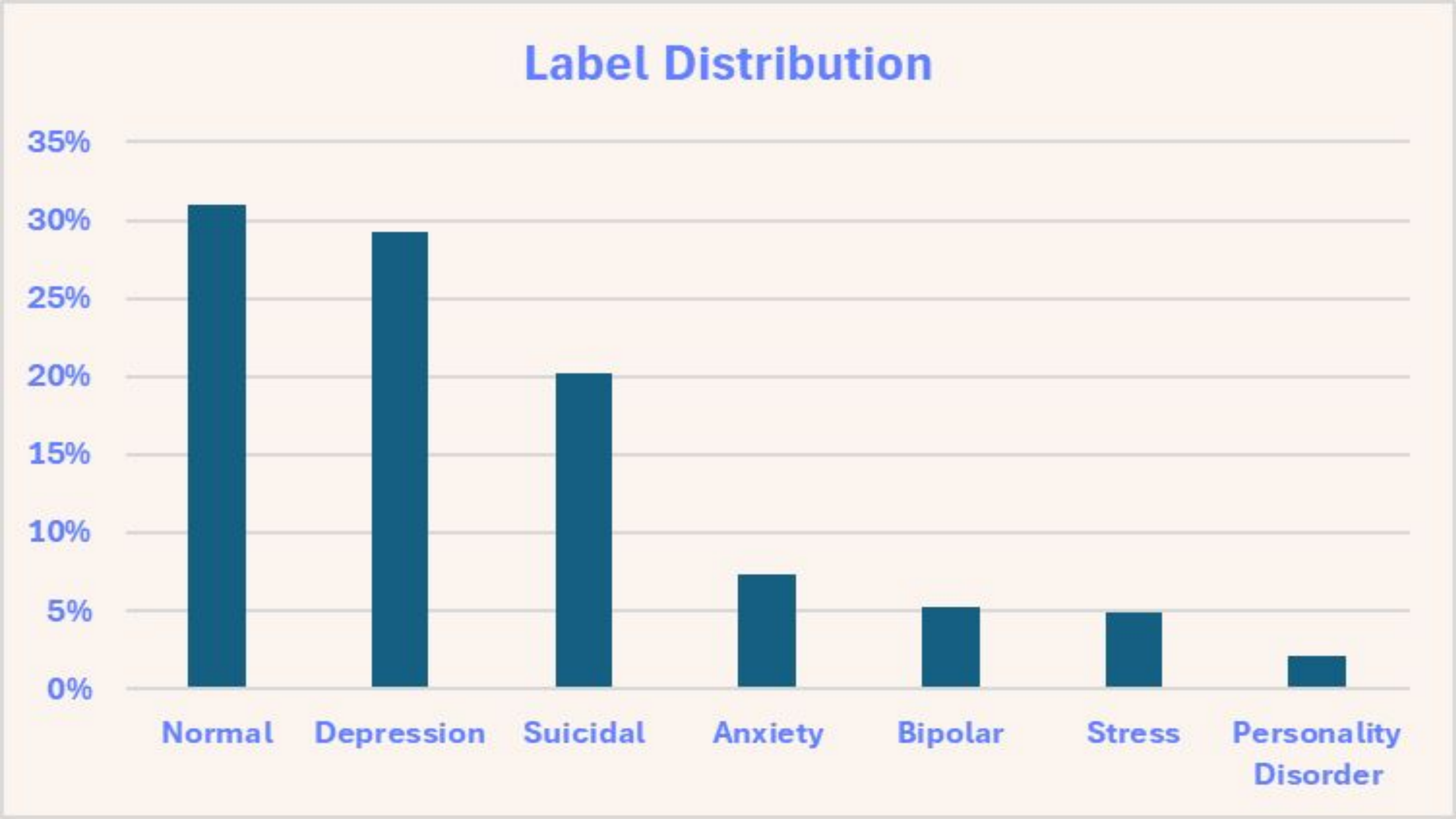
- Common in health data
- Need different metrics

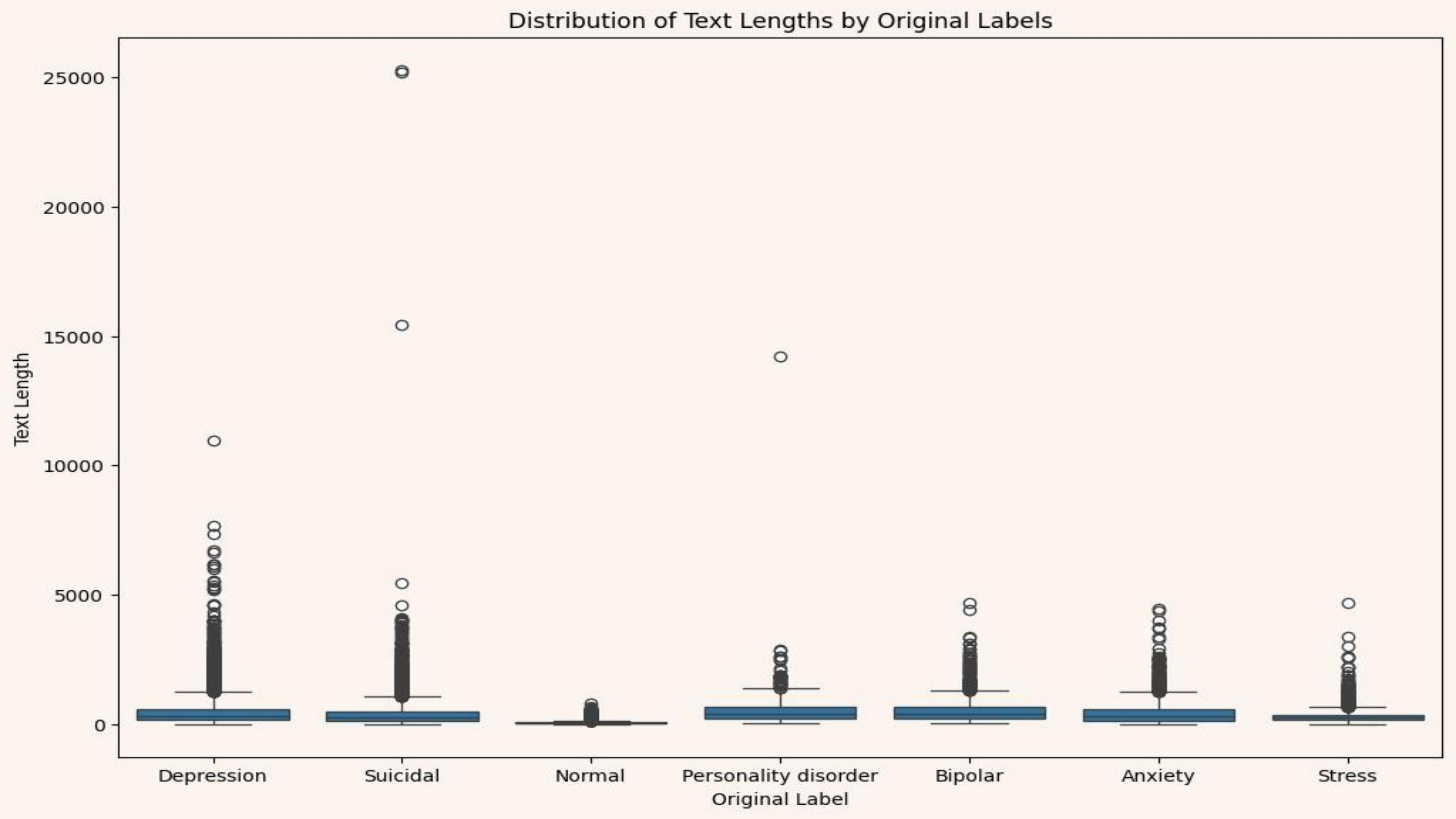
Text Length Distribution

Significant differences in “Normal” label text length

Visualizing Text

Word cloud allows us to easily and quickly visualize most frequent words in a dataset


Label Distribution
35%
30%
25%
20%
15%
10%
5%
0%
Normal
Depression
Suicidal
Anxiety
Bipolar
Stress
Personality Disorder


Distribution of Text Lengths by Original Labels
25000
20000
15000
10000
5000
0
Text Length
Depression
Suicidal
Normal
Personality disorder
Bipolar
Anxiety
Stress
Original Label

Word Cloud of Training Text Data

03

Experiments

Experiments | Multiclass Models for 7 Mental Health Statuses

Hyper-Parameters	Model 1: Bi-Directional LSTM model	Model 2: Transformer model	Model 3: Multi-hot Encoded Data	Model 4: BOW
Embedding size	128	128	NA	NA
Filter size	NA	NA	NA	NA
LSTM units	192	NA	NA	NA
Number of convolutional layers	NA	NA	NA	NA
Number of dense layers	2	11	3	NA
Number of units in each dense layer	[256, 192]	64	[256, 128]	NA
Dropout	[0.5, 0.2, 0.3]	0.5	0.2	NA
Learning Rate	1e-3	.0001	.001	NA
Number of heads	NA	4	NA	NA
Feed forward layer dimension		128	NA	NA
Weighted F1 Score	**.7416**	0.41	.7376	.7414

04

Final Models

Summary of Results

96%

Best binary model weighted F1-score

88%

Best multiclass model weighted F1-score

Most Performant Binary Model

Bag-of-words Multilayer Perceptron

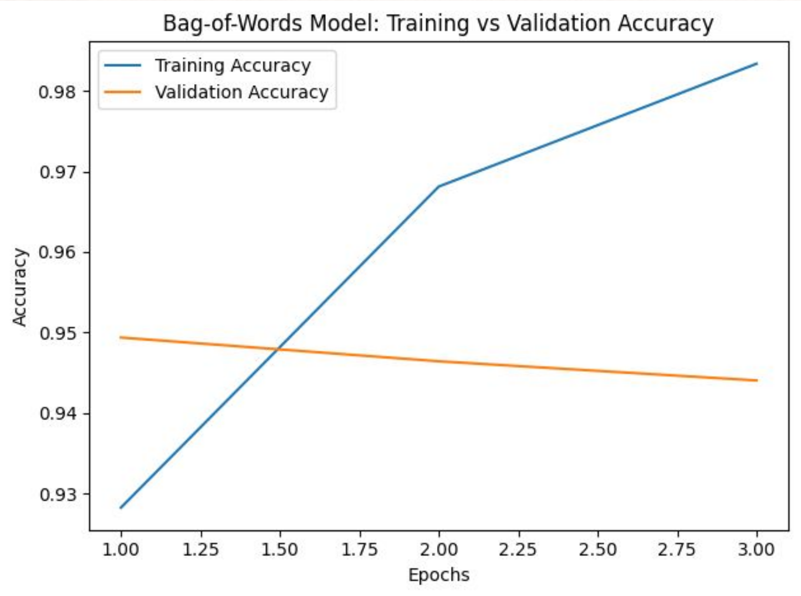
96%

F1-Score

95%

Accuracy

26% improvement over baseline accuracy



Input layer (bag-of-words vector) → Dense Layer 128 neurons activation: relu → Dropout Layer (30%) → Dense Layer 64 neurons activation: relu → Dropout Layer (30%) → Dense/output Units: 1 activation: sigmoid

Most Performant MultiClass Model

Neural Network w/Embeddings & Feature Engineering

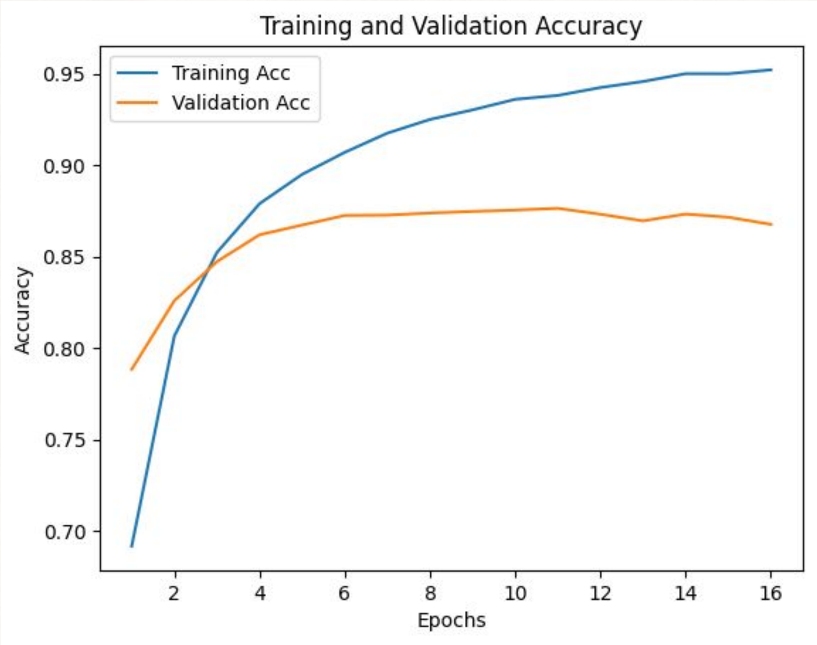
88%

F1-Score

88%

F1-Score

1. **54% improvement over baseline accuracy**
2. 5 Predicted Classes:
    a. Anxiety/Stress
    b. Depression/Suicidal
    c. Bipolar
    d. Personality Disorder
    e. None/Normal



Embedding Layer (10000, 48) → Dropout Layer (20%) → Global Average Pooling 1D → Dropout Layer (20%) → Dense/Output units: 5 activation: softmax

05

Conclusions

Conclusion

1. Key Results
   a. Successfully able to predict mental health status based on user text
      i. Bucketed Multi-Class and Binary models
2. Avenues for Future Work
   a. Pre-trained embeddings
3. Ethical considerations
   a. Training Data Bias
   b. Fairness
   c. Misdiagnosis Risk
   d. Appropriate Deployment

Contributions and GitHub Link

Aimee:Data Preprocessing (Vectorization), Modelling (Multiclass & Feature Engineering Model), Slides (Most Performant MultiClass Model)

Nura: EDA, Data Preprocessing, Modeling (Transformer Model, Bag of Words Model, Multi-hot Model, Model with Embeddings), Motivation

Wendy: Data Preprocessing (Stopwords), Modeling (Binary LSTM Model & Multiclass LSTM Model), Documentation (README)

Michael: Data Preprocessing (Lemmatization), Modelling (Bi-Directional LSTM Multiclass model), Slides (Data Preprocessing, Experiments,, Conclusion)

Abhas: EDA, Modeling (Logistic Binary Model & LSTM Binary Classification), Slides (Data Source, EDA, Conclusions)

Link to Github: https://github.com/abhaswanchu1/mids-207-final-project.git

Thank You

Appendix A1: Binary Tunings

A .csv of the configurations can be found at
https://jmp.sh/s/IGhdFUvcnjY8ktxeae4w