



## E-SAN THAILAND CODING & AI ACADEMY

โครงการวิจัยโมเดลระบบบูรณาการเรียนรู้ Coding & AI สำหรับเยาวชน

Model of Learning Ecosystem Platform integrate with Coding & AI for Youth



### โครงการย่อยที่ 6

การพัฒนาเยาวชนเพื่อเข้าสู่วิชาชีพขั้นสูงด้าน Coding & AI  
ร่วมกับ Coding Entrepreneur & Partnership: Personal AI

### AI for Detecting Users with Mental Disorders from Social media

ดร.อรุณพง วงศ์กอบลาก  
โครงการย่อยที่ 6



# อี-สาน THAILAND CODING & AI ACADEMY

## โครงการวิจัยโมเดลระบบบูรณาการเรียนรู้ที่บูรณาการ CODING & AI สำหรับเยาวชน

### Model of Learning Ecosystem Platform integrate with Coding & AI for Youth





ดี-เทคโนโลยี  
THAILAND  
CODING & AI ACADEMY

โครงการวิจัยโมเดลระบบบูรณาการเรียนรู้ที่บูรณาการ CODING & AI สำหรับเยาวชน  
Model of Learning Ecosystem Platform integrate with Coding & AI for Youth

# Outline



การพัฒนาเยาวชนเพื่อเข้าสู่วิชาชีพขั้นสูงด้าน Coding & AI ร่วมกับ Coding  
Entrepreneur & Partnership:

Personal AI

## 01 Introduction

## 02 Data Collection

## 03 Data Exploration & Preprocessing

## 04 Predictive Models

## 05 Model Evaluation

## 06 Future Work



ดี-เทคโนโลยี  
THAILAND  
CODING & AI ACADEMY

โครงการวิจัยโมเดลระบบบูรณาการเรียนรู้ที่บูรณาการ CODING & AI สำหรับเยาวชน  
Model of Learning Ecosystem Platform integrate with Coding & AI for Youth

# Outline



การพัฒนาเยาวชนเพื่อเข้าสู่วิชาชีพขั้นสูงด้าน Coding & AI ร่วมกับ Coding  
Entrepreneur & Partnership:

Personal AI

## 01 Introduction

## 02 Data Collection

## 03 Data Exploration & Preprocessing

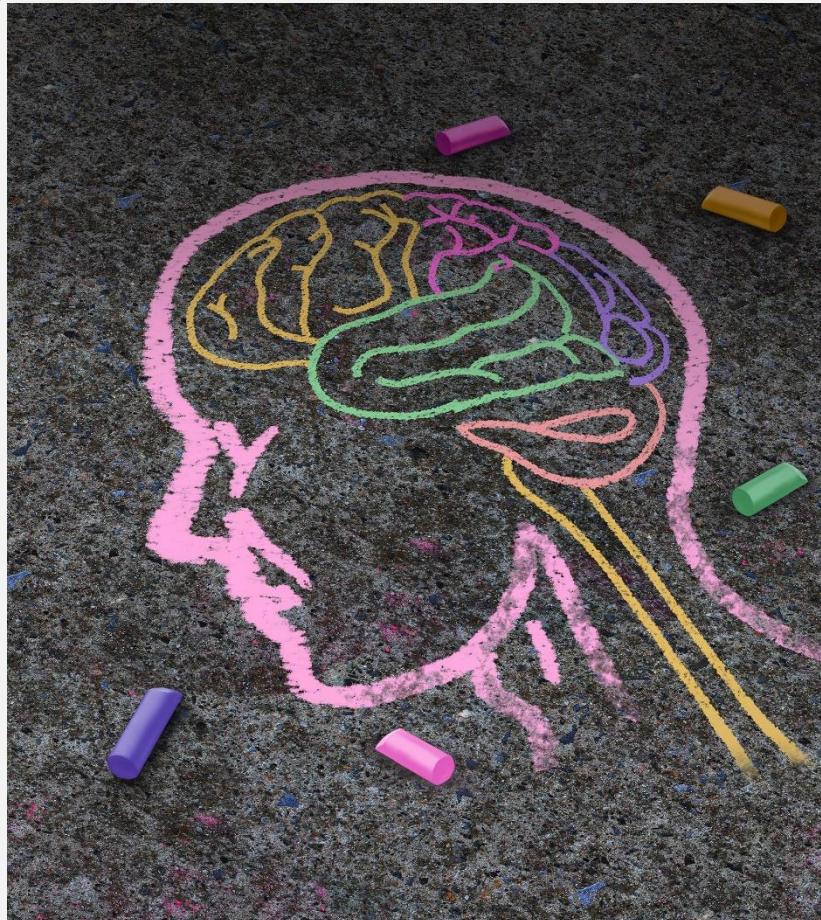
## 04 Predictive Models

## 05 Model Evaluation

## 06 Future Work



# Problem Statement



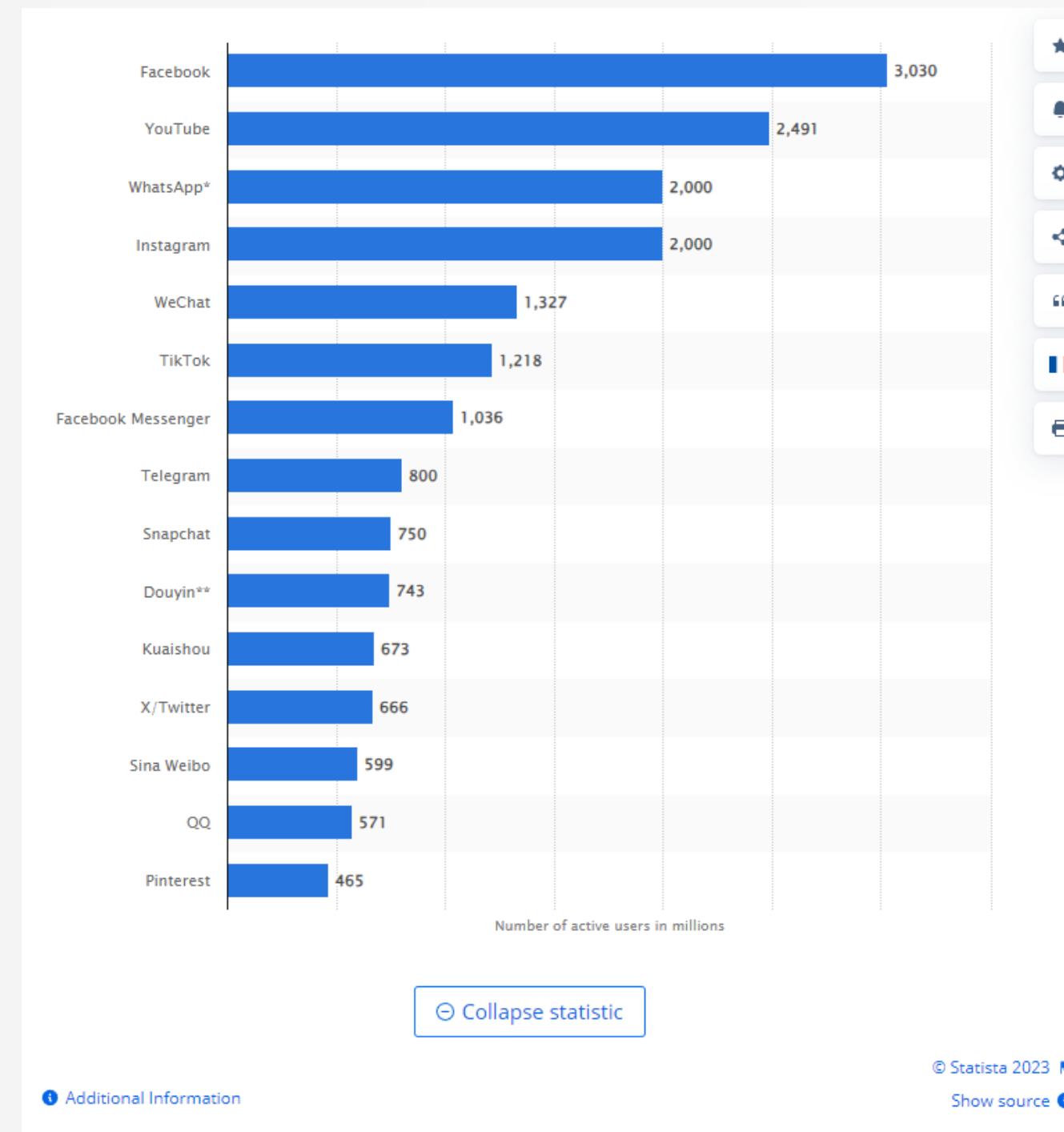
# Introduction

- **Social media** are online platforms that allow users to create, discuss, modify, and exchange content.
- Users can present their identities to others, communicate with others, form a wide variety of interest groups, and establish or maintain relationships.



**CODE COMBAT****Google****The Asia Foundation****Let's Code Thailand****DMAP  
DIGITAL MAKERS  
ASIA PACIFIC****THAI  
PROGRAMMER****E-SAN THAILAND  
CODING & AI ACADEMY**

# Introduction





CODE COMBAT

Google

The Asia Foundation

Let's Code Thailand

DMAP  
DIGITAL MAKERS ASIA PACIFIC

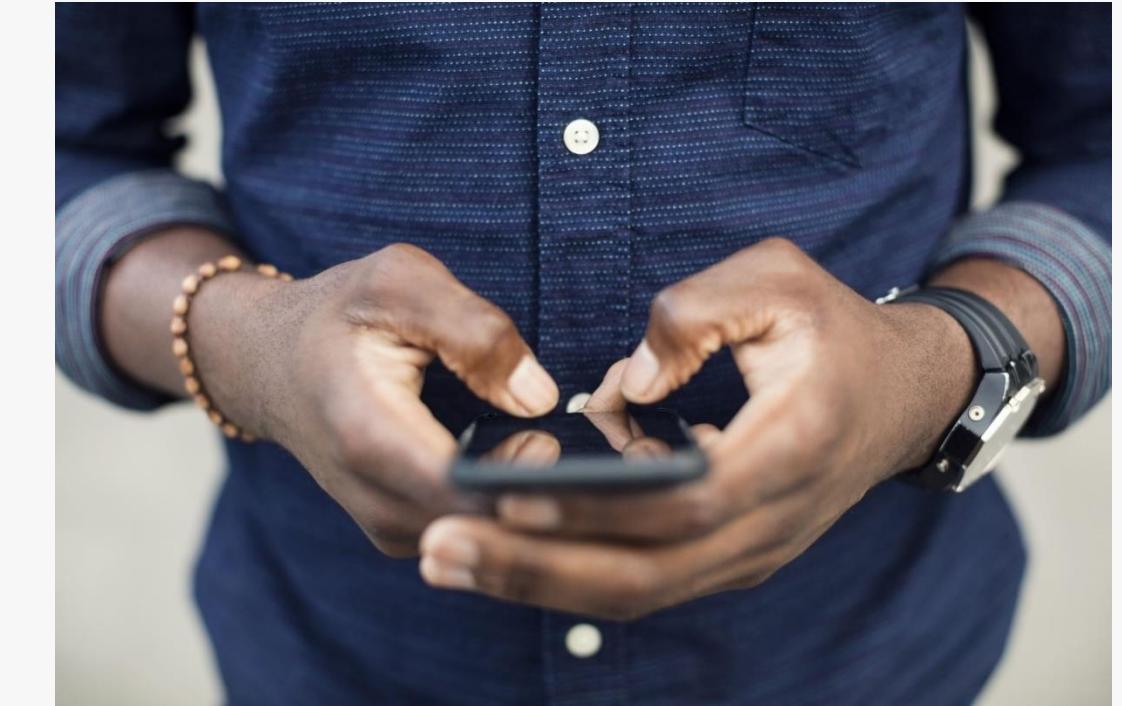
THAI  
PROGRAMMER



E-SAN THAILAND  
CODING & AI ACADEMY

# Introduction

- Immediate access
- Huge information out there
- Unfettered opinions
- Actionable insights





ดี-เทคโนโลยี  
THAILAND  
CODING & AI ACADEMY

โครงการวิจัยโมเดลระบบบูรณาการเรียนรู้ที่บูรณาการ CODING & AI สำหรับเยาวชน  
Model of Learning Ecosystem Platform integrate with Coding & AI for Youth

# Outline



การพัฒนาเยาวชนเพื่อเข้าสู่วิชาชีพขั้นสูงด้าน Coding & AI ร่วมกับ Coding  
Entrepreneur & Partnership:

Personal AI

## 01 Introduction

## 02 Data Collection

## 03 Data Exploration & Preprocessing

## 04 Predictive Models

## 05 Model Evaluation

## 06 Future Work

**CODE COMBAT****Google****E-SAN THAILAND  
CODING & AI ACADEMY**

# Types of Social Media

Type of Social Media	Description
<b>Forums</b>	A forum is an online discussion board where users can create topics to ask or exchange with other users and reply to user's topics.  <i>Example:</i> <a href="https://www.quora.com">quora.com</a> , <a href="https://www.reddit.com">reddit.com</a>
<b>Microblogs</b>	Microblogs have a word limitation of the length of posting e.g., 280 characters. Posts may contain pictures, videos, and URLs. Users can follow other accounts.  <i>Example:</i> <a href="https://www.twitter.com">twitter.com</a> , <a href="https://www.weibo.com">weibo.com</a>
<b>Products/services review</b>	Product and service review platforms are websites where users can evaluate products and write or read reviews. The sites often sell or provide product information.  <i>Example:</i> <a href="https://www.amazon.com">amazon.com</a> , <a href="https://www.yelp.com">yelp.com</a>



# Types of Social Media

Type of Social Media	Description
<b>Social networks</b>	<p>Social networking platforms allow users to create profiles and connect to others who know each other or have common interests. Users can post text, pictures, videos, and URLs.</p> <p><i>Example:</i> <a href="https://facebook.com">facebook.com</a>, <a href="https://vk.com">vk.com</a></p>
<b>Photo sharing</b>	<p>Photo sharing is a platform where users can upload, caption and share photos. Other users can comment the shared photos.</p> <p><i>Example:</i> <a href="https://flickr.com">flickr.com</a>, <a href="https://instagram.com">instagram.com</a></p>



# Social Media Matrix

	Customised Message	Broadcast Message
Profile-based	<b>Relationship</b> Allowing users to connect, reconnect, communicate, and build relationships.  (e.g., Facebook, LinkedIn)	<b>Self-Media</b> Allowing users to broadcast their updates and others to follow.  (e.g., Twitter, Weibo)
Content-based	<b>Collaboration</b> Allowing users to collaboratively find answers, advice, help, and reach consensus.  (e.g., Quora, Reddit, Yahoo! Answers)	<b>Creative outlets</b> Allowing users to share their interest, creativity, and hobbies with each other.  (e.g., YouTube, Flickr, Pinterest)



CODE COMBAT

Google

The Asia Foundation

Let's Code  
Thailand

DMAP  
DIGITAL MAKERS  
ASIA PACIFIC

THAI  
PROGRAMMER



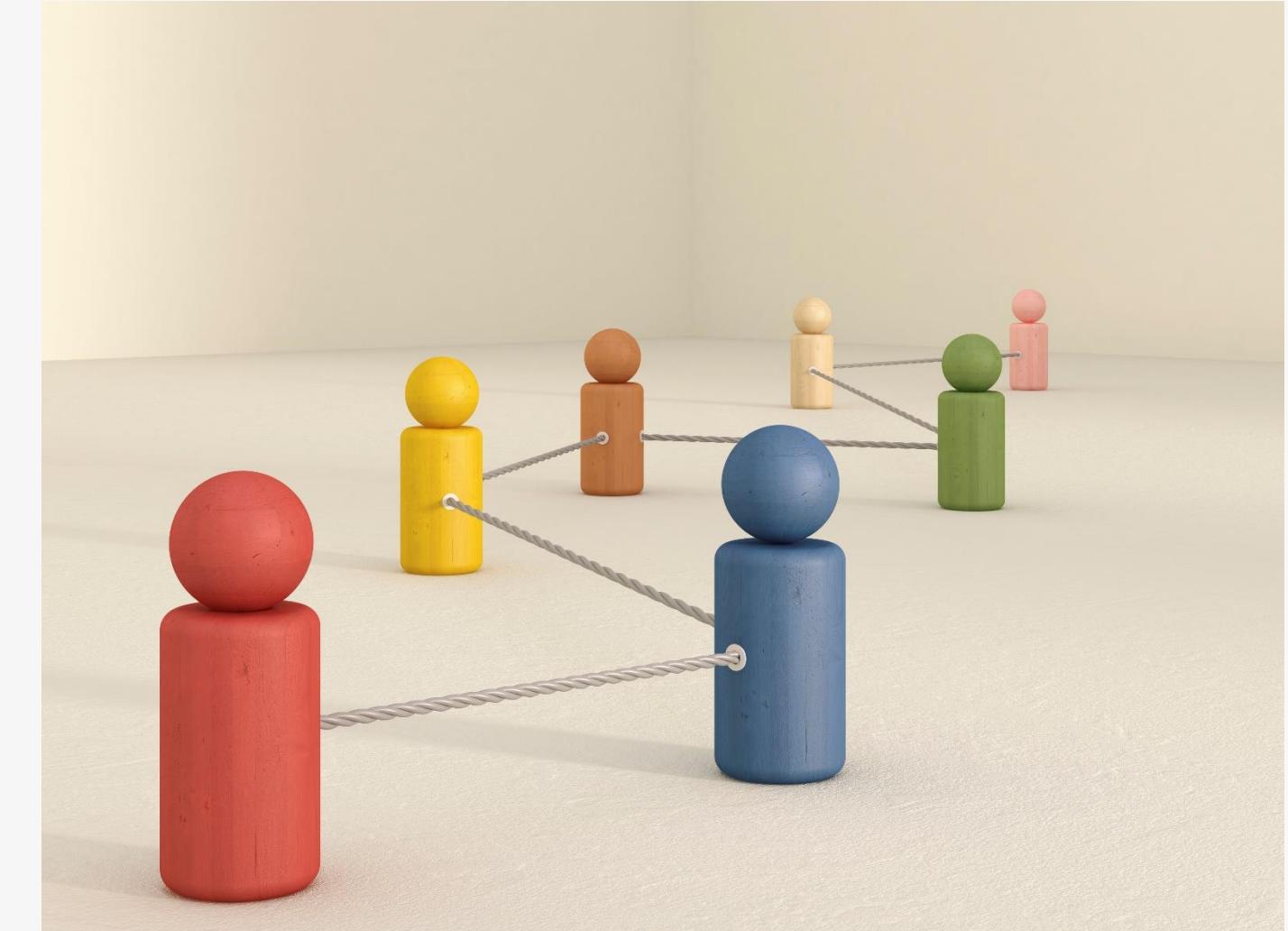
E-SAN THAILAND  
CODING & AI ACADEMY

# Data Collection

## User Data Collection



## Social Media Data Collection





CODE COMBAT

Google

The Asia Foundation

Let's Code Thailand

DMAP  
DIGITAL MAKERS  
ASIA PACIFIC

THAI  
PROGRAMMER



E-SAN THAILAND  
CODING & AI ACADEMY

# User Data Collection

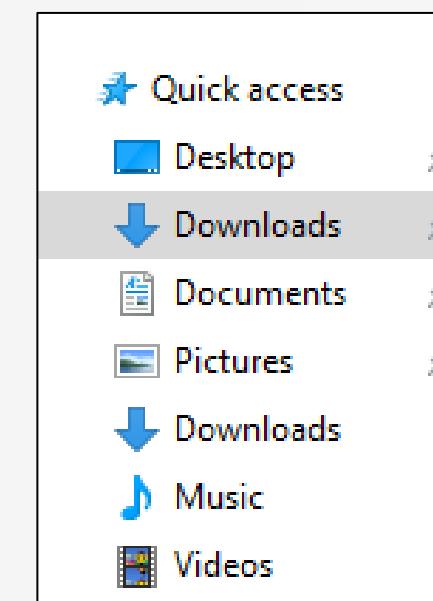
Collecting data directly from participants



Aggregating data extracted from public posts



Available Datasets



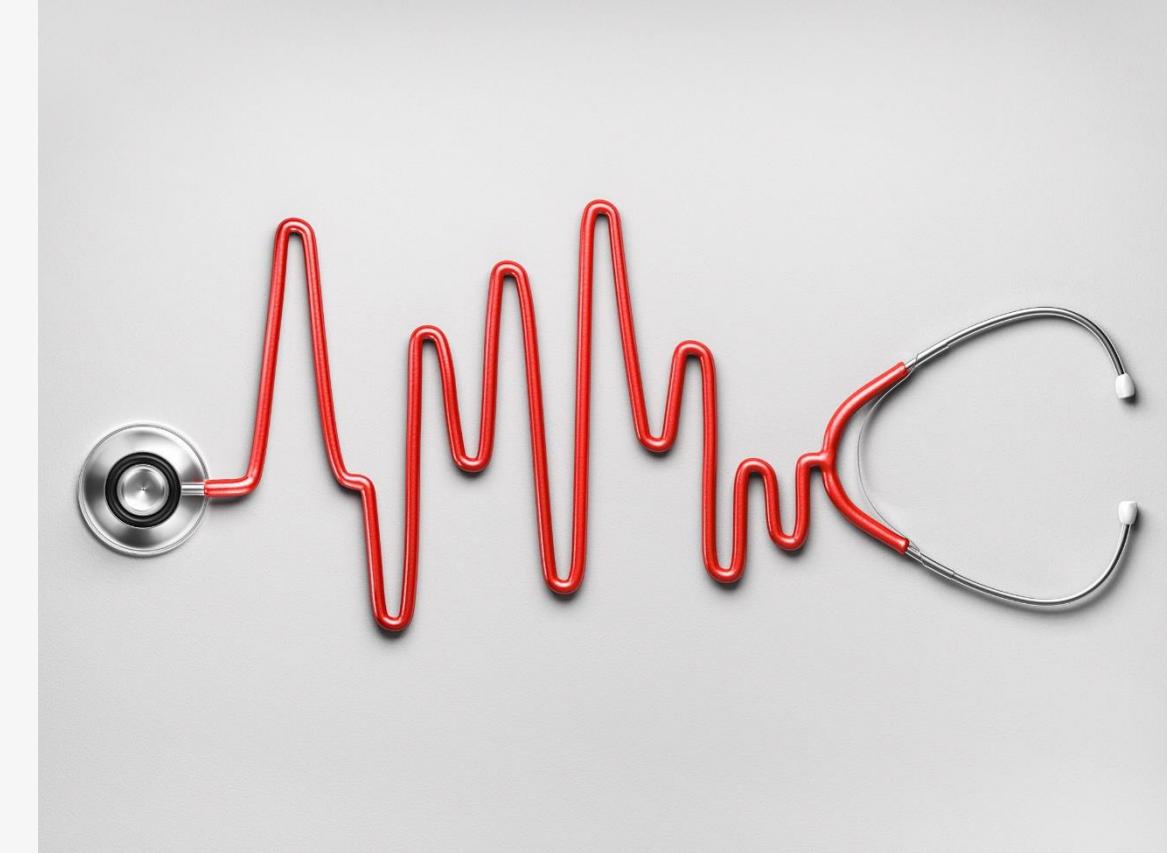


# Collecting data directly from participants

Questionnaires



EHR

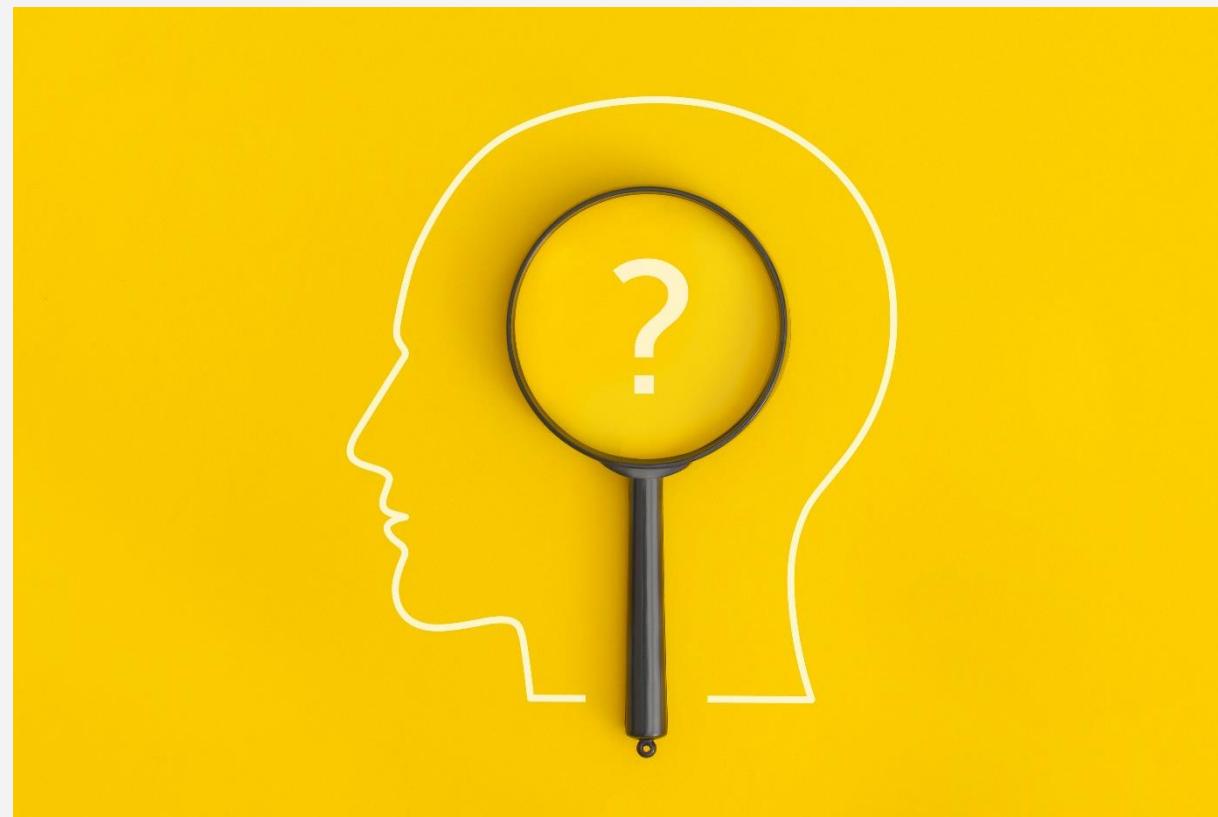




# Aggregating data extracted from public posts

Search

“I was diagnosed with [condition name]”



Annotate





ดี-เทคโนโลยี  
THAILAND  
CODING & AI ACADEMY

โครงการวิจัยโมเดลระบบบูรณาการเรียนรู้ที่บูรณาการ CODING & AI สำหรับเยาวชน  
Model of Learning Ecosystem Platform integrate with Coding & AI for Youth

# Outline



การพัฒนาเยาวชนเพื่อเข้าสู่วิชาชีพขั้นสูงด้าน Coding & AI ร่วมกับ Coding  
Entrepreneur & Partnership:

Personal AI

## 01 Introduction

## 02 Data Collection

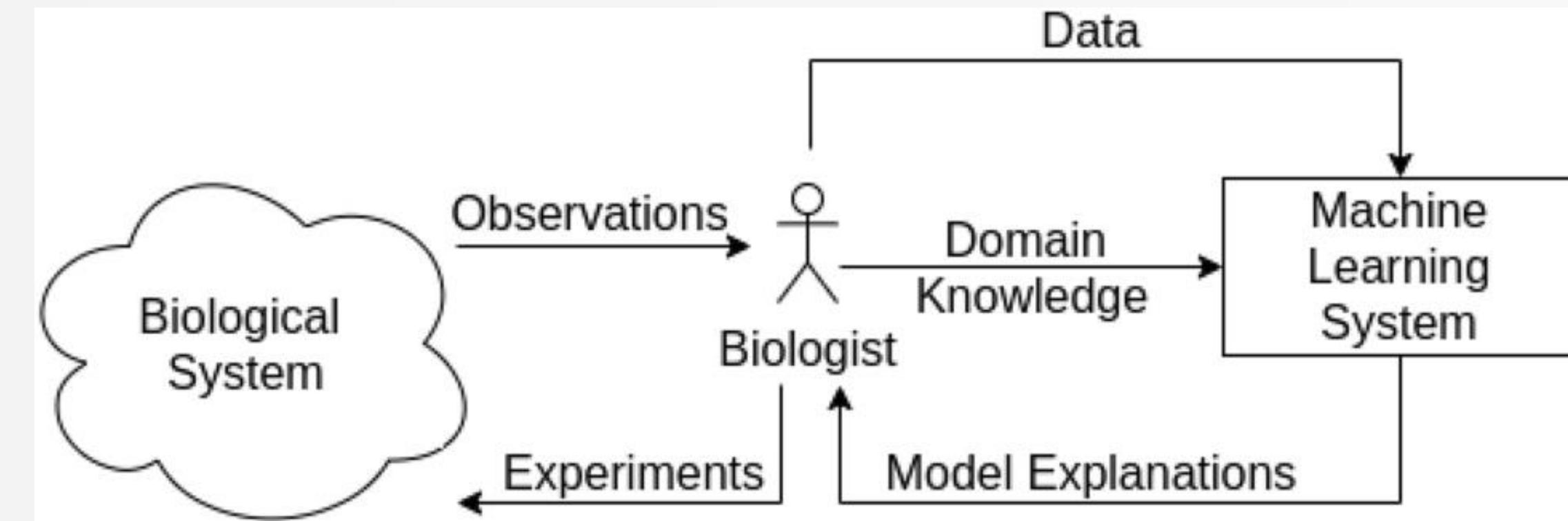
## 03 Data Exploration & Preprocessing

## 04 Predictive Models

## 05 Model Evaluation

## 06 Future Work

# Domain Knowledge





# Symptoms of depression

- Has lost interest in doing things they normally enjoy
- Seems to be feeling down or hopeless
- Has slower speech and movements or is more fidgety and restless than usual
- Feels tired or does not have much energy
- Is overeating or has lost their appetite
- Is sleeping more than usual or is not able to sleep
- Has trouble concentrating



CODE COMBAT

Google



E-SAN THAILAND  
CODING & AI ACADEMY

# Feature Extraction

```
>>> vectorizer = CountVectorizer()  
>>> corpus = [  
    'This is the first document.',  
    'This is the second second document.',  
    'And the third one.',  
    'Is this the first document?',  
>>> X = vectorizer.fit_transform(corpus)
```

```
>>> vectorizer.get_feature_names_out()  
array(['and', 'document', 'first', 'is', 'one',  
      'second', 'the',  
      'third', 'this'], ...)  
  
>>> X.toarray()  
array(  
    [[0, 1, 1, 1, 0, 0, 1, 0, 1],  
     [0, 1, 0, 1, 0, 2, 1, 0, 1],  
     [1, 0, 0, 0, 1, 0, 1, 1, 0],  
     [0, 1, 1, 1, 0, 0, 1, 0, 1]])
```

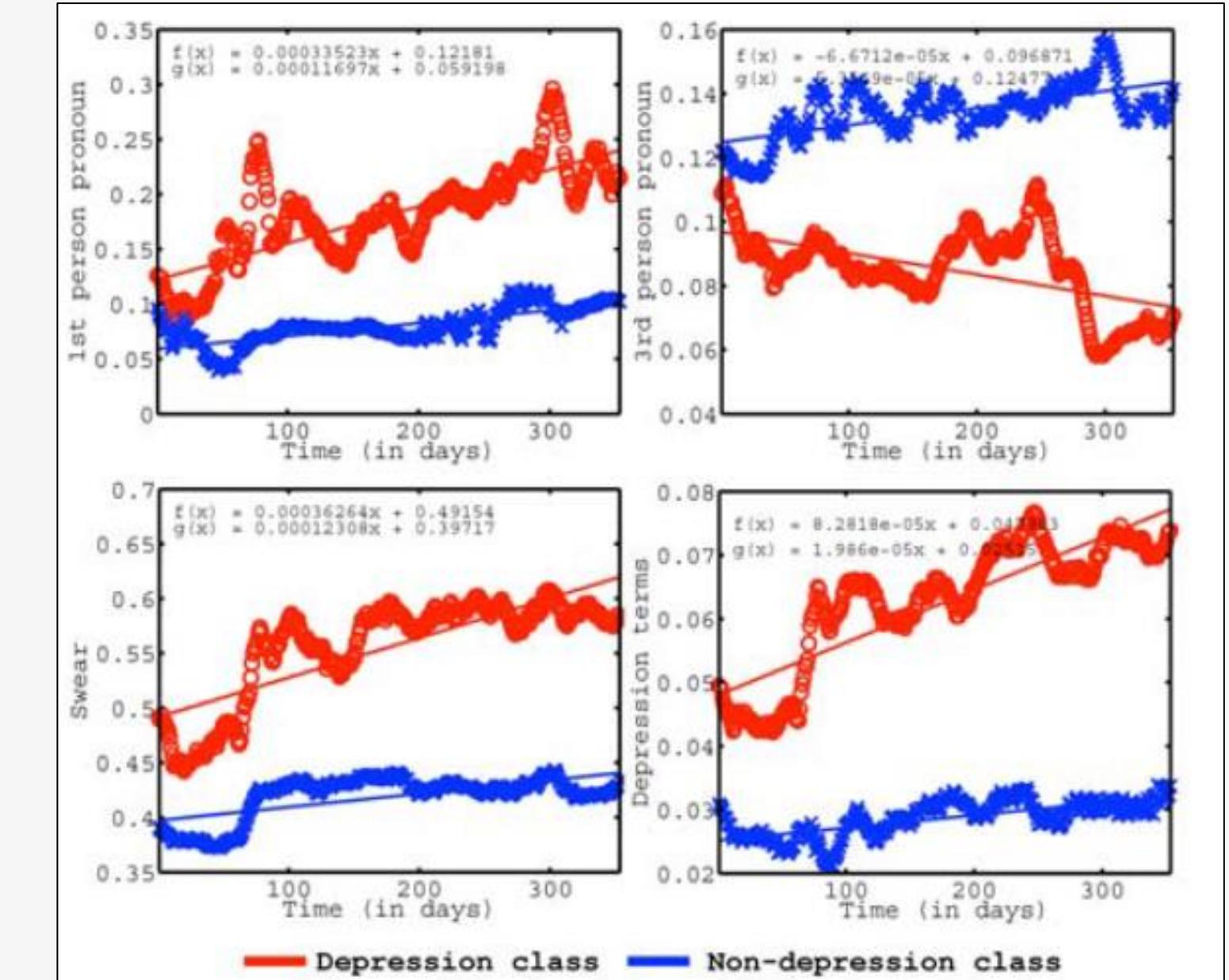
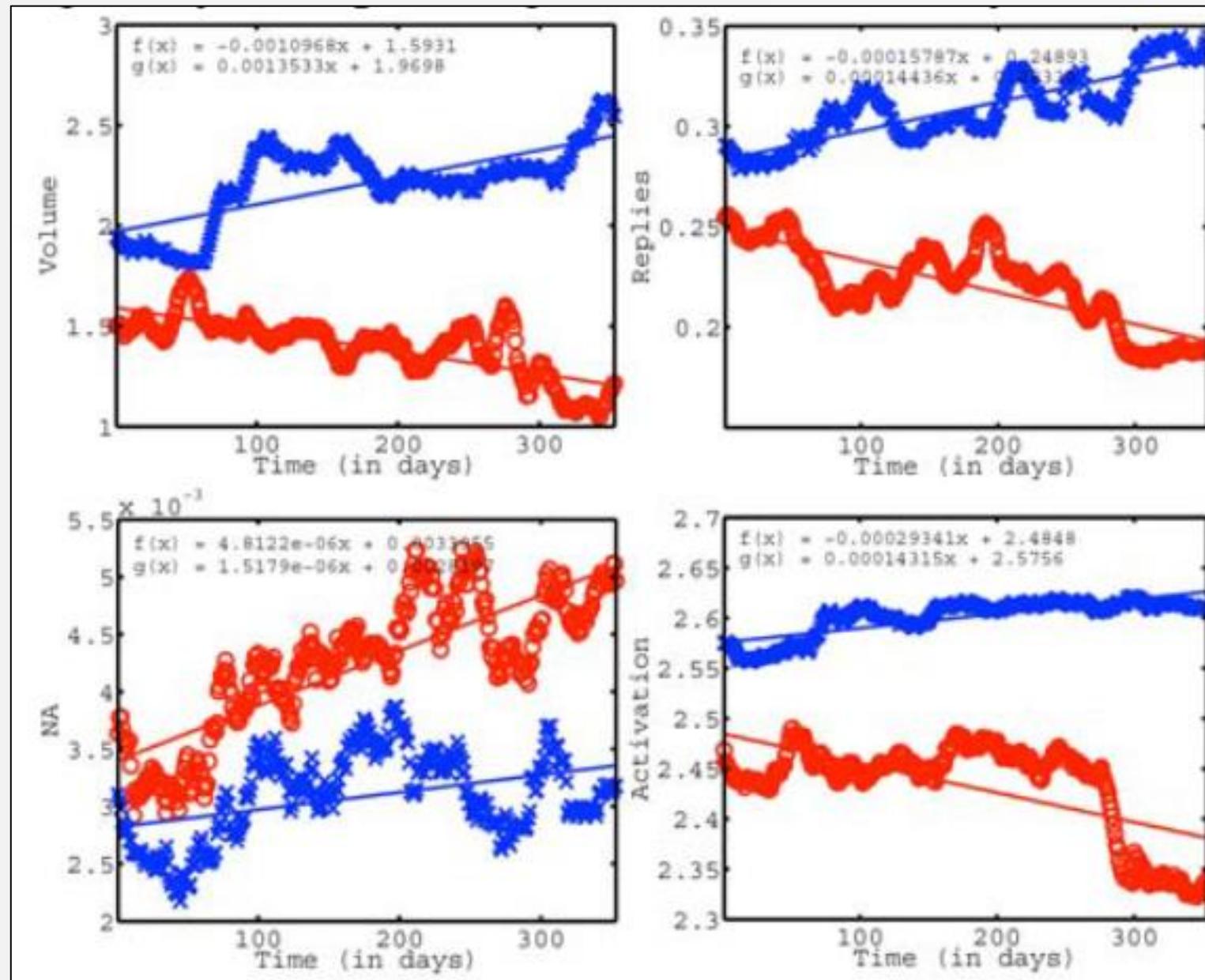


# Feature Extraction

**Table 2. LIWC-22 Language Dimensions and Reliability**

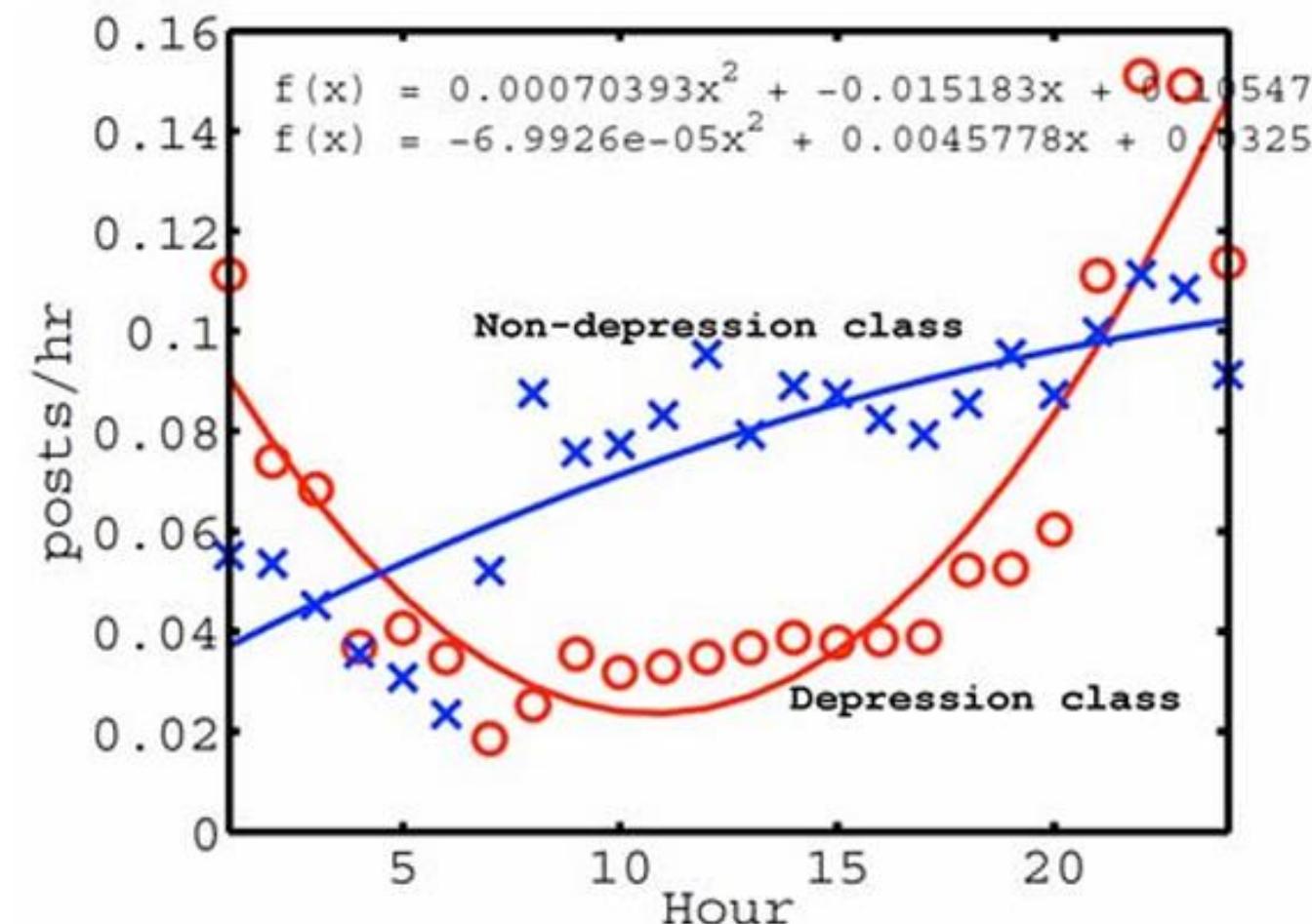
Category	Abbrev.	Description/Most frequently used exemplars	Words/ Entries in category*	Internal Consistency: Cronbach's $\alpha$	Internal Consistency: KR-20
<b>Summary Variables</b>					
Word count	WC	Total word count			
Analytical thinking	Analytic	Metric of logical, formal thinking	-	-	-
Clout	Clout	Language of leadership, status	-	-	-
Authentic	Authentic	Perceived honesty, genuineness	-	-	-
Emotional tone	Tone	Degree of positive (negative) tone	-	-	-
Words per sentence	WPS	Average words per sentence	-	-	-
Big words	BigWords	Percent words 7 letters or longer	-	-	-
Dictionary words	Dic	Percent words captured by LIWC	-	-	-
<b>Linguistic Dimensions</b>					
Total function words	function	the, to, and, I	499/1443	0.36	1.00
Total pronouns	pronoun	I, you, that, it	74/286	0.43	0.97
Personal pronouns	ppron	I, you, my, me	42/221	0.24	0.95
1st person singular	i	I, me, my, myself	6/74	0.49	0.85
1st person plural	we	we, our, us, lets	7/17	0.43	0.78
2nd person	you	you, your, u, yourself	14/59	0.37	0.82
3rd person singular	shehe	he, she, her, his	8/30	0.58	0.83
3rd person plural	they	they, their, them, themsel*	7/20	0.36	0.69
Impersonal pronouns	ipron	that, it, this, what	32/68	0.43	0.91
Determiners	det	the, at, that, my	97/293	-0.19	0.95
Articles	article	a, an, the, a lot	3/103	0.12	0.61
Numbers	number	one, two, first, once	44/61	0.57	0.87

# Data Exploration





# Data Exploration





## ดี-เทคโนโลยี THAILAND CODING & AI ACADEMY

โครงการวิจัยโมเดลระบบบูรณาการเรียนรู้ที่บูรณาการ CODING & AI สำหรับเยาวชน  
Model of Learning Ecosystem Platform integrate with Coding & AI for Youth

# Outline



การพัฒนาเยาวชนเพื่อเข้าสู่วิชาชีพขั้นสูงด้าน Coding & AI ร่วมกับ Coding  
Entrepreneur & Partnership:

Personal AI

## 01 Introduction

## 02 Data Collection

## 03 Data Exploration & Preprocessing

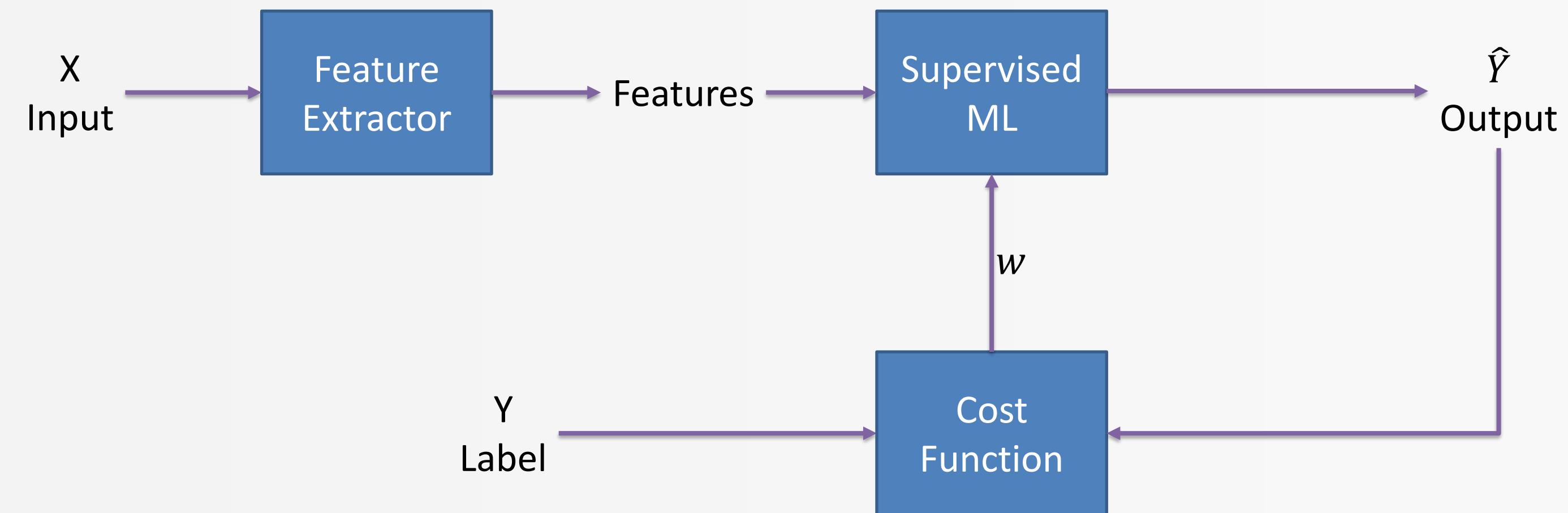
## 04 Predictive Modelling

## 05 Model Evaluation

## 06 Future Work



# NLP with Supervised Learning





CODE COMBAT

Google

The Asia Foundation

Let's Code Thailand

DMAP  
DIGITAL MAKERS ASIA PACIFIC

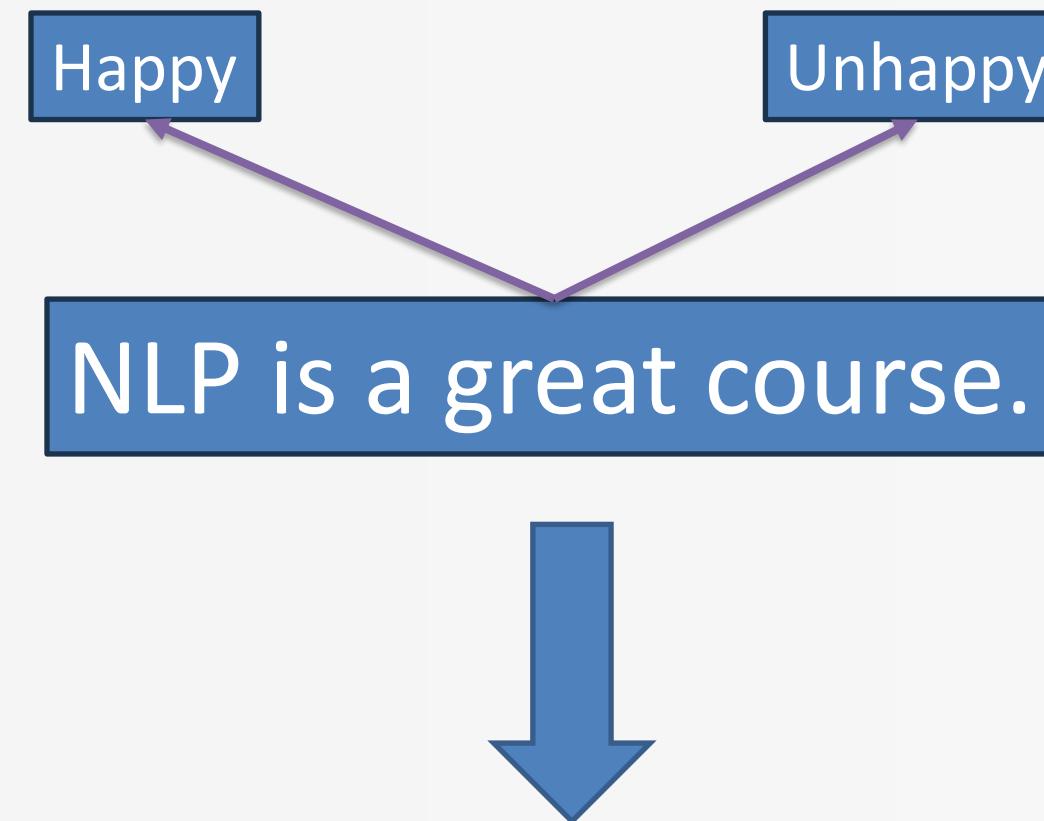
THAI  
PROGRAMMER



E-SAN THAILAND  
CODING & AI ACADEMY

# NLP with Supervised Learning

Word Count

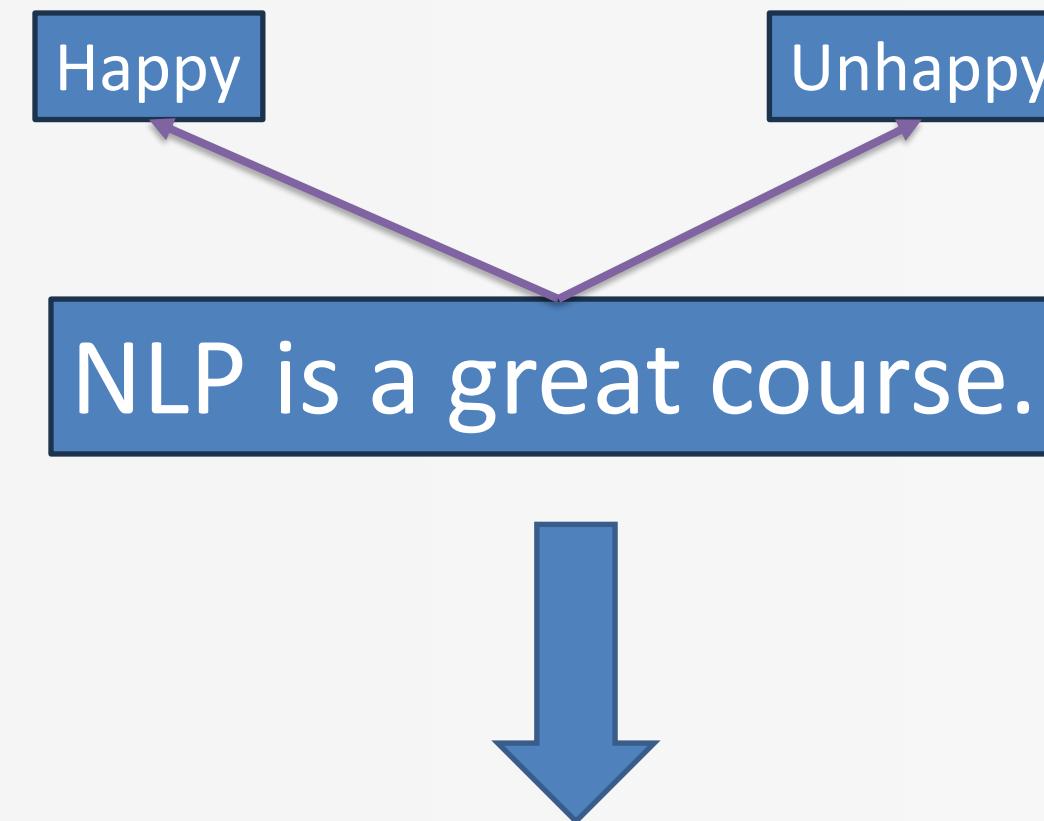


[1, 1, 1, 1, 1]



# NLP with Supervised Learning

Positive/Negative Words

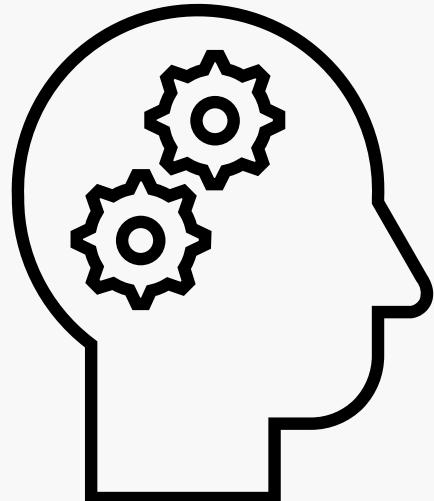


[1, 0]

**CODE COMBAT****Google****E-SAN THAILAND CODING & AI ACADEMY**

# Predictive Modelling

Label	date	post	n_chars	n_sents	n_words	sent_neg	sent_neu	sent_pos	mic_stress	olation_tot
mentalhealth	01/01/2018	Any idea w	885	22	244	0.114	0.784	0.102	1	0
mentalhealth	01/01/2018	Advice,	1091	24	260	0.14	0.682	0.177	1	0
mentalhealth	01/01/2018	Can	241	7	58	0.039	0.961	0	0	0
mentalhealth	01/01/2018	I heard my	1897	42	503	0.128	0.7609999	0.111	6	0
mentalhealth	01/01/2018	From the	2201	39	580	0.145	0.737	0.118	3	0
mentalhealth	01/01/2018	2018 is no	863	20	224	0.152	0.63	0.218	0	0
mentalhealth	01/01/2018	How do I t	974	21	260	0.04	0.937	0.023	1	1
mentalhealth	01/01/2018	Help Hi	3463	28	822	0.092	0.852	0.056	6	0
mentalhealth	01/01/2018	I don't kno	383	6	100	0.168	0.753	0.079	0	0
mentalhealth	01/01/2018	I need a	1963	30	512	0.123	0.7879999	0.089	7	0
mentalhealth	01/01/2018	Curing	698	13	153	0.134	0.76	0.106	0	0
mentalhealth	01/01/2018	What wou	684	10	157	0.109	0.765	0.126	1	0
mentalhealth	01/01/2018	I	642	15	167	0.21	0.711	0.08	0	0
mentalhealth	01/01/2018	I'm so	753	20	199	0.111	0.845	0.044	0	0
mentalhealth	01/01/2018	Really ner	364	8	96	0.085	0.825	0.09	1	0





ดี-เทคโนโลยี  
THAILAND  
CODING & AI ACADEMY

โครงการวิจัยโมเดลระบบบูรณาการเรียนรู้ที่บูรณาการ CODING & AI สำหรับเยาวชน  
Model of Learning Ecosystem Platform integrate with Coding & AI for Youth

# Outline



การพัฒนาเยาวชนเพื่อเข้าสู่วิชาชีพขั้นสูงด้าน Coding & AI ร่วมกับ Coding  
Entrepreneur & Partnership:

Personal AI

## 01 Introduction

## 02 Data Collection

## 03 Data Exploration & Preprocessing

## 04 Predictive Models

## 05 Model Evaluation

## 06 Future Work

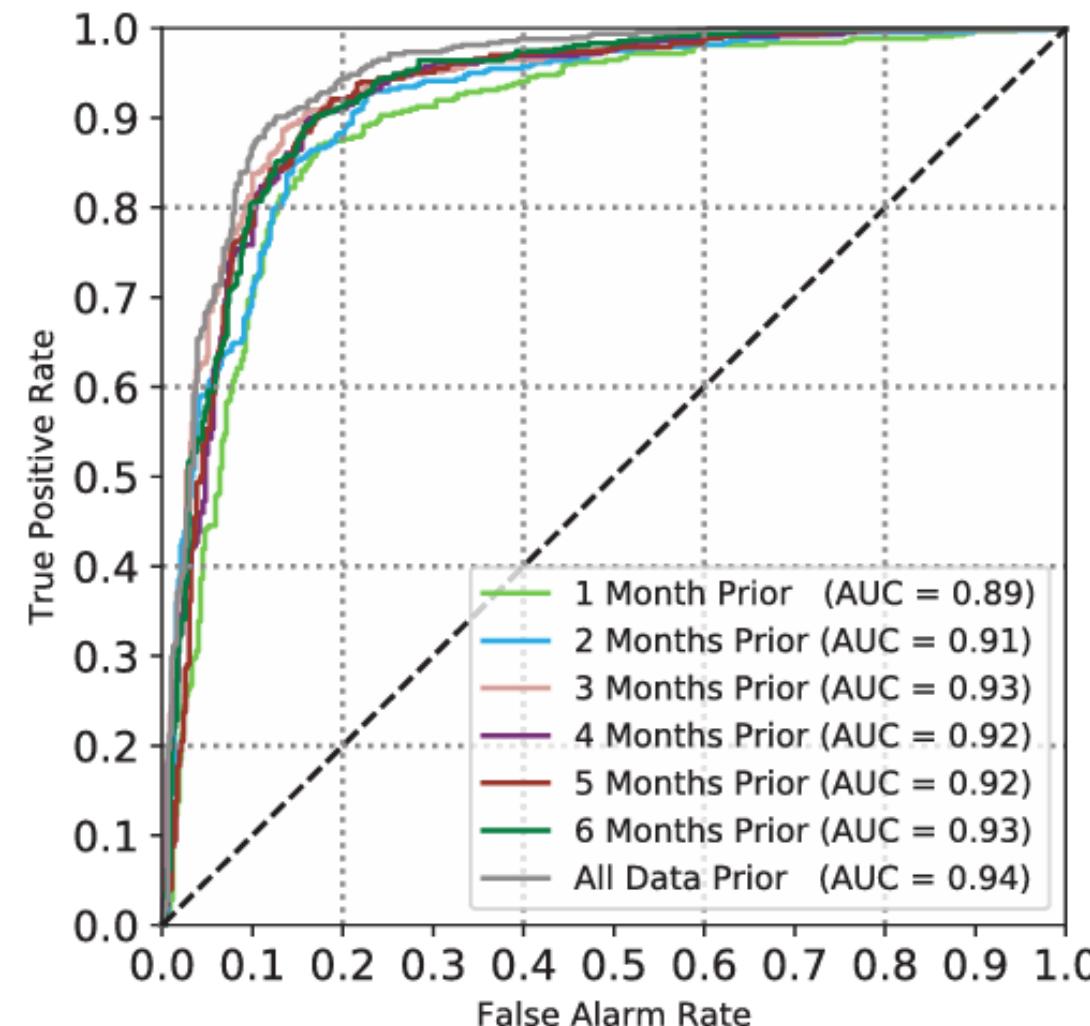


# Evaluation

	precision	recall	acc. (+ve)	acc. (mean)
engagement	0.542	0.439	53.212%	55.328%
ego-network	0.627	0.495	58.375%	61.246%
emotion	0.642	0.523	61.249%	64.325%
linguist. style	0.683	0.576	65.124%	68.415%
dep. language	0.655	0.592	66.256%	69.244%
demographics	0.452	0.406	47.914%	51.323%
all features	0.705	0.614	68.247%	71.209%
dim. reduced	<b>0.742</b>	<b>0.629</b>	<b>70.351%</b>	<b>72.384%</b>

Table 6. Performance metrics in depression prediction in posts using various models. Third column shows the mean accuracy of predicting the positive class.

# Evaluation



**Figure 6.** ROC curves for models separating users prior to a suicide attempt from their matched controls. The green line only uses data for the month prior to the suicide attempt to make the classification (30 to 0 days prior), the blue line uses data from 2 months prior (60 to 0 days prior), and so on. The black line indicates performance using all of the data available for that user prior to their attempt. ROC indicates receiver operating characteristic.



## ดี-เทคโนโลยี THAILAND CODING & AI ACADEMY

โครงการวิจัยโมเดลระบบบูรณาการเรียนรู้ที่บูรณาการ CODING & AI สำหรับเยาวชน  
Model of Learning Ecosystem Platform integrate with Coding & AI for Youth

# Outline



การพัฒนาเยาวชนเพื่อเข้าสู่วิชาชีพขั้นสูงด้าน Coding & AI ร่วมกับ Coding  
Entrepreneur & Partnership:

Personal AI

## 01 Introduction

## 02 Data Collection

## 03 Data Exploration & Preprocessing

## 04 Predictive Models

## 05 Model Evaluation

## 06 Future Work



# Future Work

