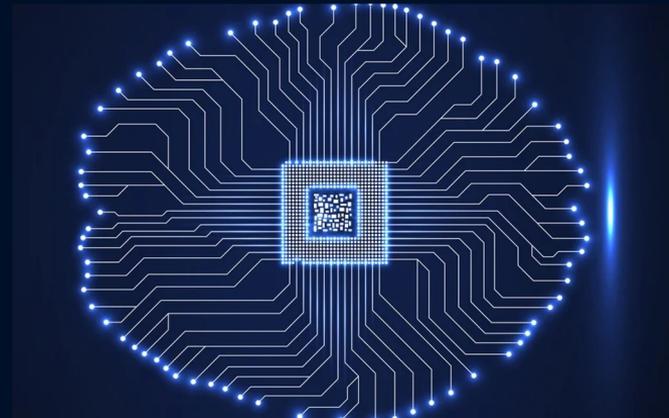
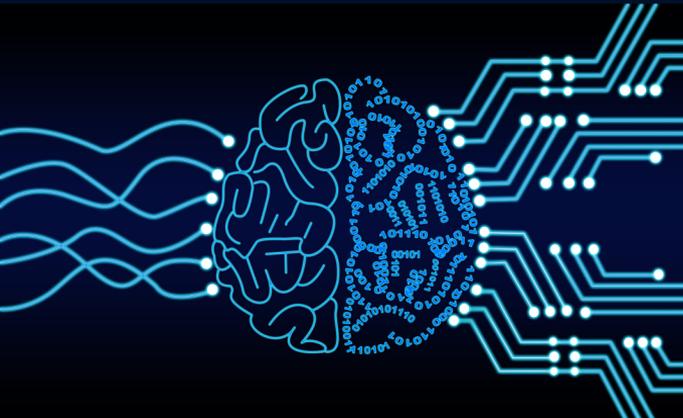
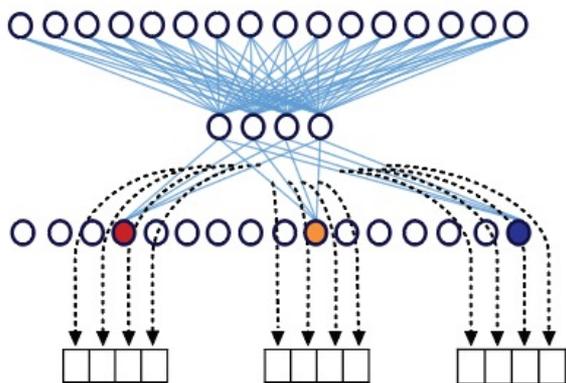


Natural Language Processing in Neuropsychiatry

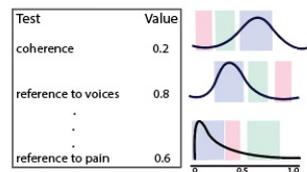
Neguine Rezaii, MD
Instructor of Neurology
Harvard Medical School
March 28, 2022



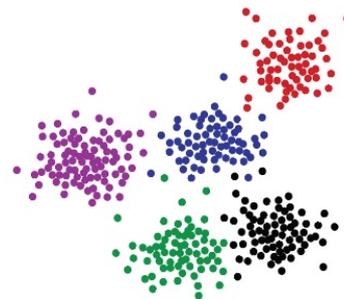
Feeding NLP models with patients' direct language



Yes, I believe I have been feeling this way for a long time... you know some sort of fuzziness in my head which is really difficult to describe... it comes and goes though



- normal
- delirium due to:
 - intoxication
 - infection



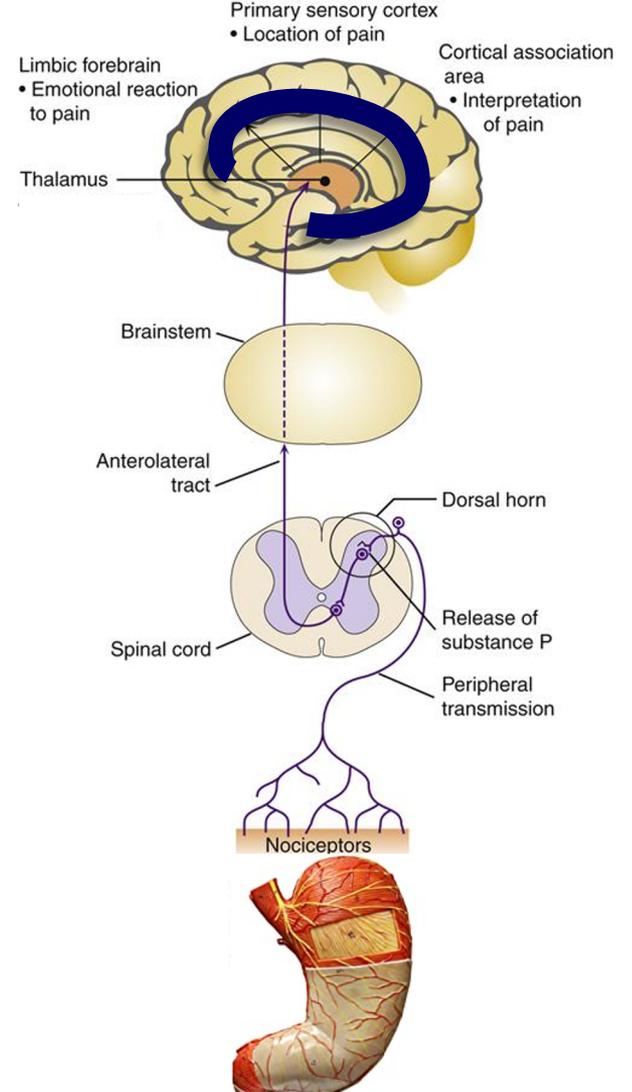
- normal
- delirium due to:
 - intoxication
 - infection
 - surgery
 - emotional distress



Why language?

The language network has access to the information contained in a wide range of organs, including the physically distant ones

Direct message, e.g. having stomach ache





Why language?

Indirect message, e.g. when we experience a vague discomfort

Use of a particular phrase, word order, conversational maxim or even acoustic feature can provide clues to a person's physiological or pathological state

Our everyday language contains information about the functional status of our mind and body

Language also contains clues about the more subliminal information

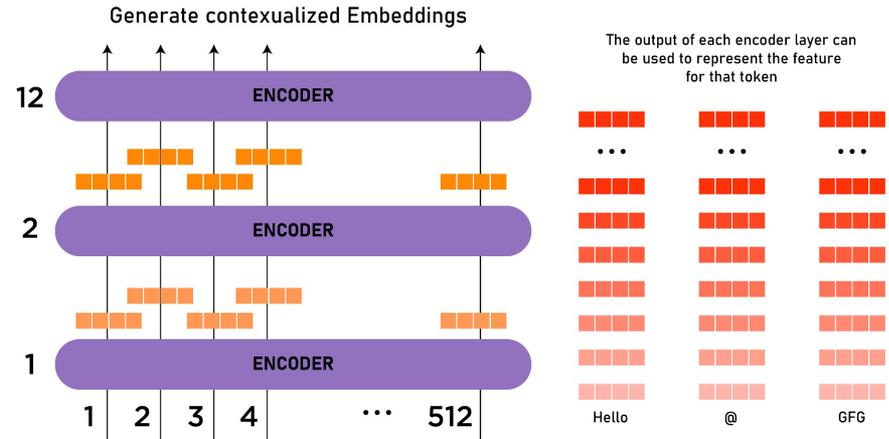
Psychologists from Yale University showed that people judged others to be more generous and caring if they had just held a warm cup of coffee and less so if they had held an iced coffee (Williams and Bargh, 2008, Science)

From an individual's language, knowledge about their mental and physiological states, and even environmental factors such as temperature can be extracted

Reverse engineering



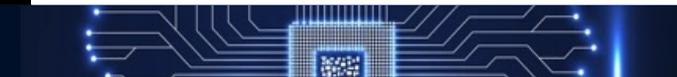
- The signals are often too subtle to be heard with the *naked ear*
- This type of reverse engineering requires large amount of language samples
- NLP models acquire their knowledge of syntax, semantics and pragmatics from large amounts of text and store that knowledge in layers of artificial neural networks.
- The resulting knowledge can be harnessed to address multiple long-standing problems in psychiatry





1. Making the subjective objective

- Clinicians learn to pick up on the key indicators of illness in people's communication, including their spoken language, gestures, facial expressions, silences, eye contact and posture
- Typically, these aspects of communication are evaluated in a qualitative manner, making them vulnerable to bias and inconsistency across clinicians
- Tools from NLP can be used to help make these impressions more objective



Measuring Semantic Density

I think I do have strong feelings for politics.

So you cannot... you cannot... well actually I'm supreme, nobody will not try me.

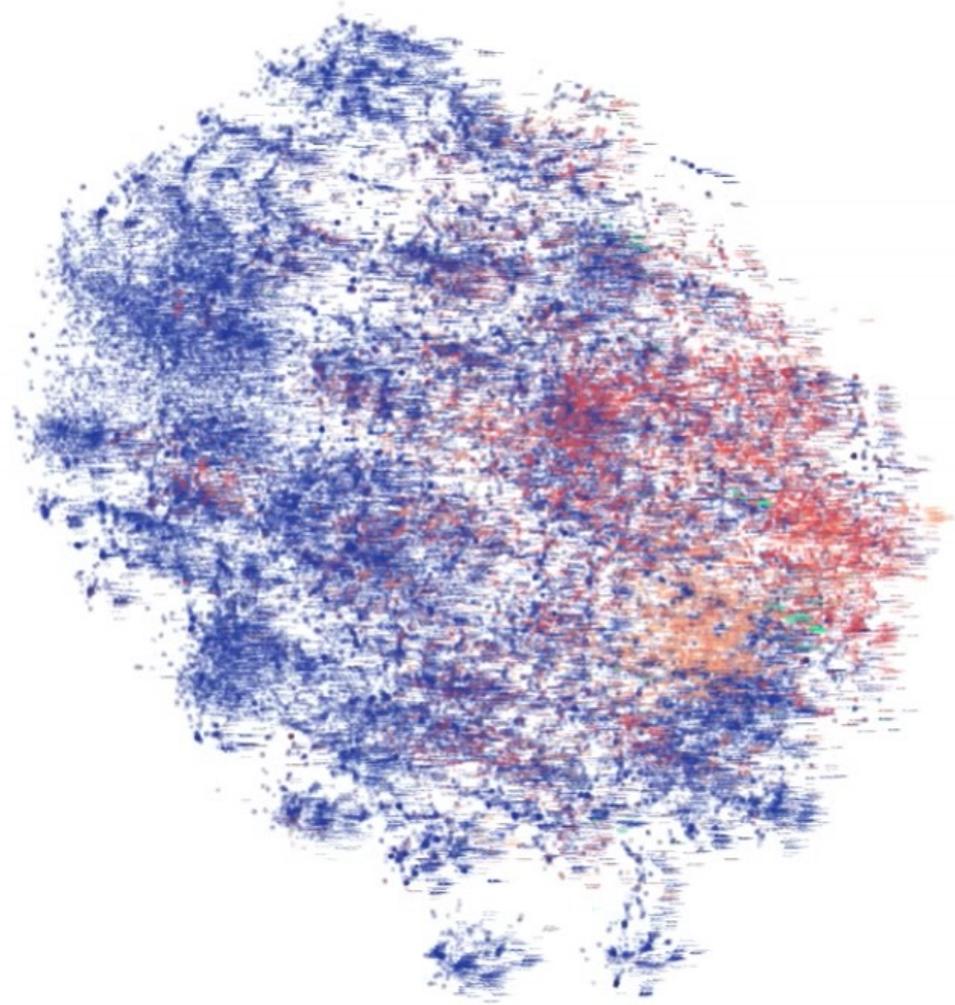


Word2Vec



25 years of text from the *New York Times* which included 42,833,581 sentences

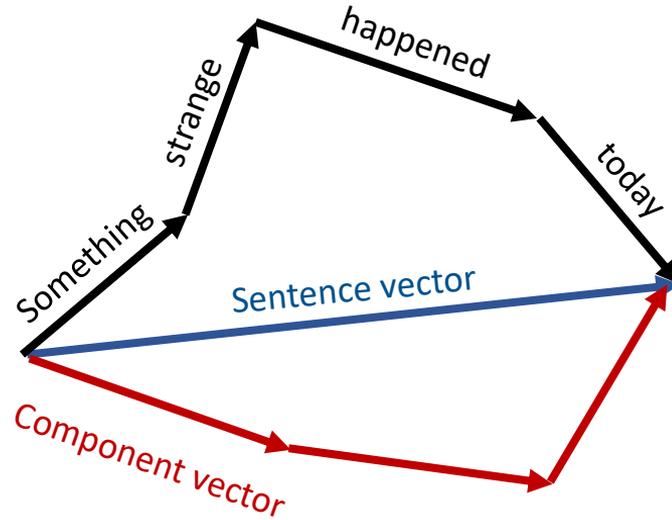




Lucky =

| | | | | | | | |
|-------------------|--------------------|-------------------|-------------------|--------------------|-------------------|-------------------|-------------------|
| 0.0206761534995 | -0.164801059576 | -0.00687682216626 | -0.159940788401 | 0.0716044693507 | 0.0468098247222 | 0.0265225229492 | -0.0261137376182 |
| 0.0478563369576 | 0.0261489273484 | 0.00927473153724 | 0.0758457413215 | -0.0226959646302 | -0.0290799426944 | 0.0756958451366 | 0.025970663373 |
| -0.0463597150636 | -0.023197848554 | -0.00476262155251 | 0.0442813314494 | -0.0462907618405 | 0.0305954674591 | -0.0105308829958 | 0.0683922168921 |
| 0.0435341750482 | -0.065097876134 | 0.124472114842 | 0.0215575598483 | -0.000630625393998 | -0.0483998171592 | -0.052548712453 | -0.121162158523 |
| 0.134346793306 | 0.12613265716 | 0.0272671092389 | -0.0147121931908 | -0.0320263456019 | 0.0119911300627 | 0.0161306826716 | -0.036870853571 |
| 0.00398759868493 | 0.112405351855 | 0.0373012069629 | 0.0769541178645 | 0.00807312072076 | -0.0828252007688 | -0.0503926215838 | -0.110072138135 |
| -0.0522452880436 | -0.00891801923966 | 0.0289439520737 | -0.0314336646472 | -0.0763161267018 | -0.0278959267487 | -0.0877344463697 | -0.0370046669149 |
| -0.0686444108835 | 0.102875904382 | 0.0319004267248 | -0.00592105163179 | -0.00866912500271 | -0.0842025958871 | -0.0149116261256 | -0.0460688383224 |
| 0.0381943608047 | -0.0190290962472 | 0.0330253244325 | 0.0174948867433 | -0.10869129479 | -0.0381816387828 | -0.0675295891993 | -0.0702947232828 |
| 0.0477174463536 | -0.045437007808 | 0.0252212119195 | -0.101560082134 | 0.0995290651873 | -0.0206176459987 | 0.0368115777044 | 0.0783352371045 |
| 0.201550696013 | 0.0803636847911 | 0.171386353075 | 0.076697841017 | 0.0353496816413 | 0.0232224139157 | 0.0649613147052 | -0.0628097408779 |
| 0.0223424296786 | 0.00147089838051 | 0.0632625962146 | 0.0114128552918 | 0.0852616903914 | 0.0777856961772 | -0.0741532629675 | -0.0302200151423 |
| -0.0122023480753 | -0.000494802659409 | -0.0177053323507 | 0.0308049674694 | -0.059250133107 | 0.0250187754388 | 0.0203028543544 | 0.0797951612438 |
| -0.0389806941962 | 0.1142207989914 | -0.103573774251 | -0.0157357060605 | -0.0919566221399 | -0.031025247608 | -0.0450285485102 | 0.0783713044695 |
| 0.146505920738 | 0.0595562210341 | -0.0559000786942 | -0.0219388306837 | 0.0886290632462 | 0.035342435172 | -0.00589243633864 | 0.0763029353569 |
| 0.0335694018574 | 0.0287890910521 | -0.046372160373 | 0.147246391511 | 0.038002650595 | -0.094306493379 | 0.030362952978 | 0.0358670525864 |
| 0.0508138156459 | -0.0850359091019 | 0.042962134585 | 0.0574105929335 | -0.0809763370454 | 0.095490295178 | 0.0377114192575 | -0.037788382221 |
| -0.0874655308306 | -0.0236713685531 | 0.114602068563 | 0.145021545109 | -0.0748451487468 | -0.0879603492279 | -0.0441050518486 | 0.010247851252 |
| -0.0418237712481 | -0.0133805314098 | -0.063484769743 | 0.114763749621 | -0.0483609384969 | 0.00901073648055 | -0.061063477789 | 0.0163561179314 |
| 0.130271294418 | -0.0761822539465 | 0.193865328252 | -0.0477681822022 | 0.0964599804675 | 0.0469661588193 | 0.0356776292854 | -0.141982437584 |
| -0.00401430353083 | 0.0315571721057 | -0.116809291993 | 0.0942023339842 | 0.0901725603807 | -0.0233941649612 | -0.0119009667563 | 0.0040190050557 |
| -0.0989502666744 | 0.149602545706 | 0.0414707697058 | -0.172683213502 | 0.102449782057 | 0.0577096780627 | 0.0625211017212 | -0.0102679830111 |
| 0.121783622213 | -0.09778237511 | -0.00266692010908 | 0.0560649035657 | -0.00165309660074 | -0.00906110854195 | -0.0363302789466 | -0.0170452932856 |
| -0.0952330183697 | 0.135259183548 | 0.0313292937723 | 0.0637082605472 | 0.0457181956837 | -0.0134338067338 | 0.0660711266083 | 0.139916577265s2 |
| 0.060006478728 | -0.0480226458434 | -0.0447684576504 | -0.0570445036989 | 0.0188789708159 | 0.162940078291 | -0.0503810955578 | 0.0607210918253 |
| 0.0510784248022 | -0.130687578893 | -0.0365469152105 | -0.0921784396303 | 0.0727587187478 | 0.0670373922538 | 0.0671313027151 | -0.00741577968679 |

Something strange happened today.



Semantic Density = Number of component vectors / Number of words

$$3 / 4 = 0.75$$

Well, I think I do have strong feelings for politics.

1.0



I mean it seems like as much as we want to believe we matter in universe, I don't think there is really anything that suggests that we do.

0.93



Sometimes I get nervous and I get worried about time and I can't look at the clock.

0.62



So you cannot... you cannot... well actually I'm supreme,
nobody will not try me.

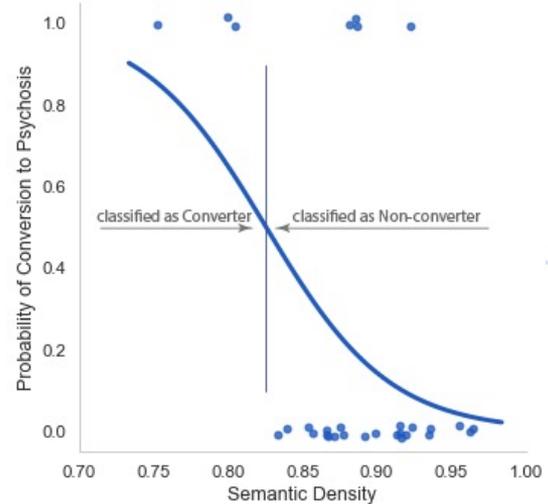
0.14





Predicting conversion to schizophrenia based on semantic density

Semantic Density predicted conversion to psychosis with an accuracy of 80%





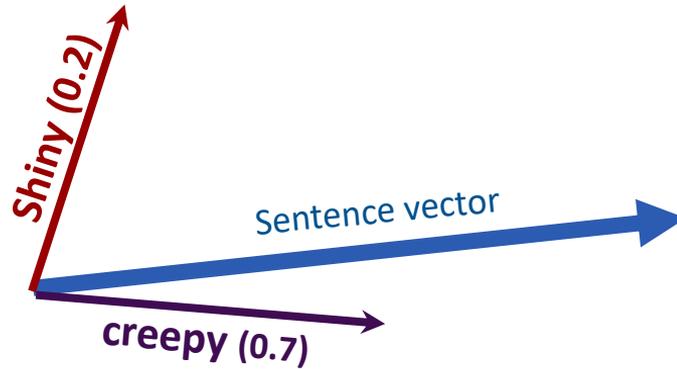
2. Discovering symptoms that are easily missed

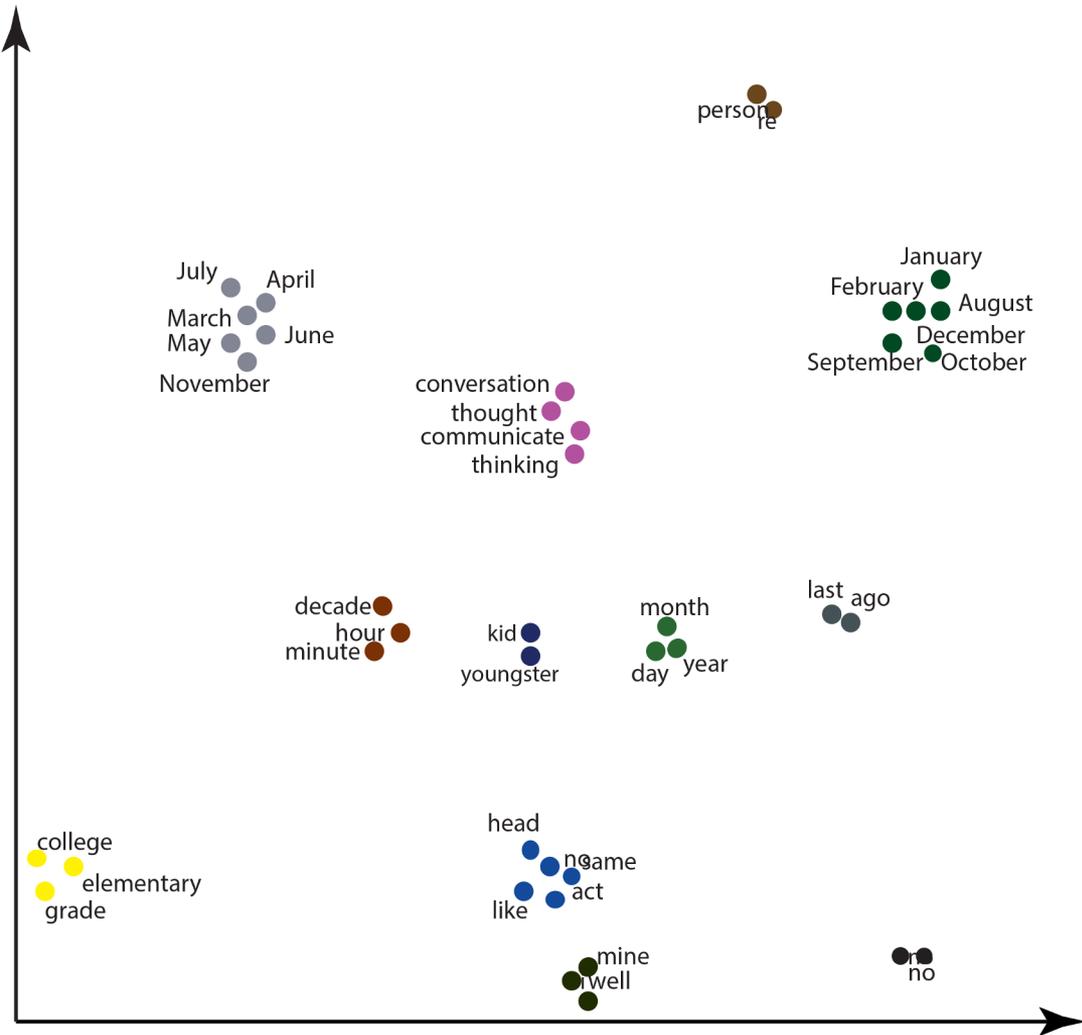
Methods from NLP can go beyond what is overtly stated by revealing hidden dimensions of meaning, thereby addressing a long-sought goal in psychiatric disciplines such as psychoanalysis



Extracting latent content

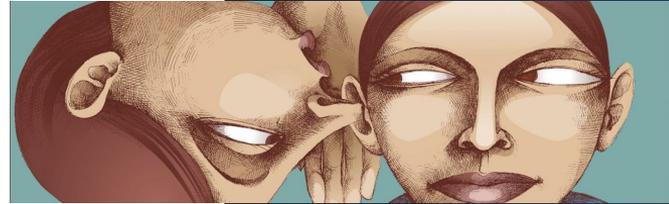
Something strange happened today.





Patient 1) “...You know I talk to myself but I don’t ... I don’t know if it is me. I mean if I talk to myself in the mirror you know I’m talking to me. But how can I have a conversation with myself? I say stuff in my head as if I am talking to me and it’s funny and I laugh like I didn’t know that I was going to say that...”

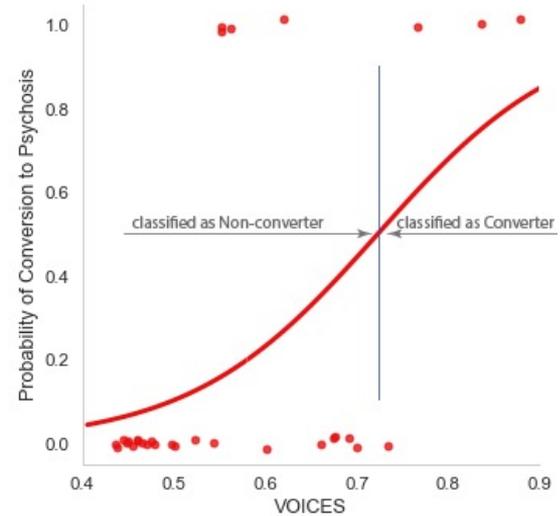
Patient 2) “I would hear something that sound like a plane engine or like a really... you know... a really far off motor. It never went away entirely. It’s gone a lot more in the past couple of months since Christmas. It just sounds like that... it sounds like a little flame or a cellular... a digital motor.”





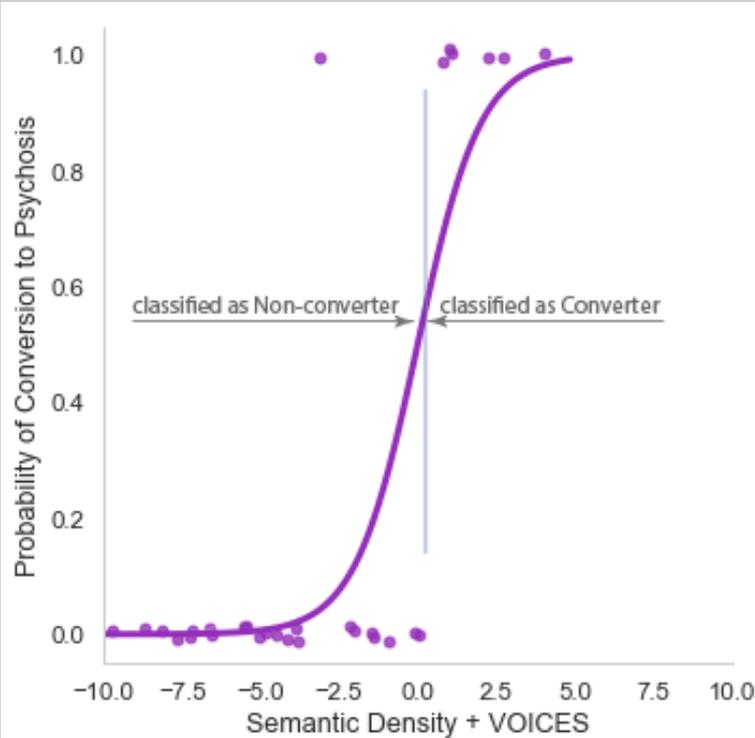
Predicting conversion to schizophrenia based on reference to voice

Reference to voice predicted conversion to psychosis with an accuracy of 83.3%



Using both semantic density and reference to voice increased the accuracy of prediction of conversion to 93%

$$\text{Conversion} = -57.254 \times \text{SEMANTIC DENSITY} + 20.483 \times \text{VOICES} + 35.828$$



3. Classification of diseases

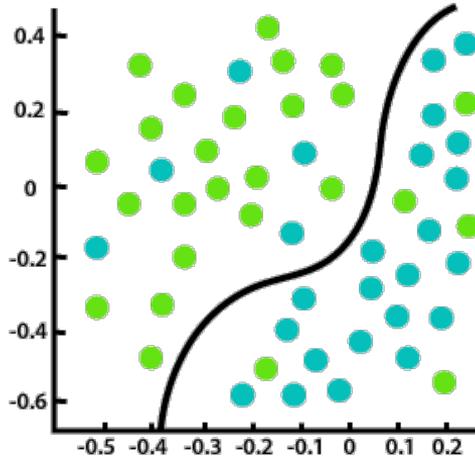
Theory-driven approaches

Sometimes the disease categories in psychiatry are clear, but the symptoms determining the assignment of a patient to a particular disease category are not

For example, suicide is an unambiguous disease outcome, but predicting the likely occurrence of suicide has been a challenge

However, with the help of NLP and machine learning tools as master pattern recognizers, occurrence of suicide can be predicted by an individual's word choices and vocal characteristics.

Crucially, the models for predicting a particular disease integrate hundreds of linguistic and acoustic features, likely more than can be simultaneously considered by a human observer



3. Classification of diseases

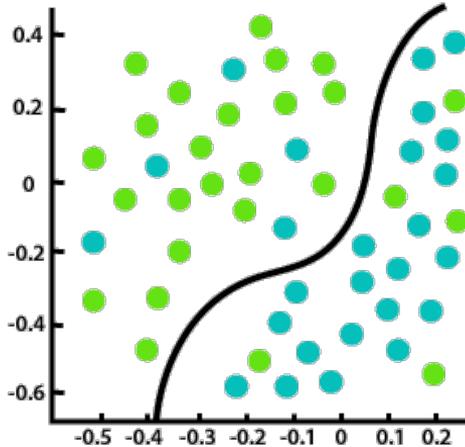
Data-driven approaches

At other times, the symptoms are clear, but the disease categories are not well understood.

NLP methods can be used to help discover and define natural categories of disease by measuring similarities in language production across individuals and determining the number of ways these individuals tend to cluster in terms of this data.

For example, natural clusters in primary progressive aphasia have been identified on the basis of the language features of syntactic complexity, hesitation rate and naming

This approach is of particular value in fields such as neuropsychiatry where unifying disease classification theories are still evolving



Conclusion

Language is replete with layers of information about a person's mental and physiological states

This information often goes beyond the direct message that the individual communicates

Advances in NLP models now make it possible to extract the latent indicators of mental health from language to improve the diagnosis and prediction of neuropsychiatric disorders



Thank You!



@NeguineR

This drawing depicts various styles of written language including cuneiform and cave painting

