

 This site uses cookies for analytics, personalized content and ads. By continuing to browse this site, you agree to this use.

[Learn more \(https://go.microsoft.com/fwlink/?linkid=845480\)](https://go.microsoft.com/fwlink/?linkid=845480)

 Microsoft  
(<http://msdn.microsoft.com/>)

| Developer

Search



Sign in ([https://blogs.msdn.microsoft.com/wp-](https://blogs.msdn.microsoft.com/wp-login.php?aadso_action=login)

[login.php?aadso\\_action=login](https://blogs.msdn.microsoft.com/wp-login.php?aadso_action=login))

# Humanizing Artificial Intelligence (AI) with Deep Learning

Rate this article ★★★★★

 Puneet Agrawal (BING) (<https://social.msdn.microsoft.com/profile/Puneet+Agrawal+%28BING%29/>) October 6, 2017

2 (<https://blogs.msdn.microsoft.com/msind/2017/10/06/humanizing-artificial-intelligence-ai-with-deep-learning/#comments>)

 Share 140

68

0

## Subscribe Blog via Email

Subscribe to this blog and receive notifications of new posts by email.

Email Address

Subscribe!

Unsubscribe

## Tags

machine learning  
(<https://blogs.msdn.microsoft.com/msind/tag/machine-learning/>)

artificial intelligence  
(<https://blogs.msdn.microsoft.com/msind/tag/artificial-intelligence/>)

deep learning  
(<https://blogs.msdn.microsoft.com/msind/tag/deep-learning/>)

Microsoft  
(<https://blogs.msdn.microsoft.com/msind/tag/microsoft/>)

AI  
(<https://blogs.msdn.microsoft.com/msind/tag/ai/>)

DL  
(<https://blogs.msdn.microsoft.com/msind/tag/dl/>)

Digital India  
(<https://blogs.msdn.microsoft.com/msind/tag/digital-india/>)

Chatbots  
(<https://blogs.msdn.microsoft.com/msind/tag/chatbots/>)

SMS  
(<https://blogs.msdn.microsoft.com/msind/tag/sms/>)

Kaizala  
(<https://blogs.msdn.microsoft.com/msind/tag/kaizala/>)

DNN  
(<https://blogs.msdn.microsoft.com/msind/tag/dnn/>)

are conveyed through expressions and tone of voice.

At Microsoft, we are working to create human-like AI, Ruuh (<http://facebook.com/ruuh>) and on this journey detecting user emotions is a critical piece. So, I teamed up with Microsoft researchers Umang Gupta and Radhakrishnan Srikanth to take on this challenge. Also, Ankush Chatterjee, an intern from IIT Kharagpur, joined us taking on his first machine learning research assignment!

Here's a look at how we arrived at a novel approach to detect emotions in textual conversations:

### The challenge of creating emotional Artificial Intelligence (AI)

Detecting emotions accurately is always a challenge, even for human beings. Participants in a conversation often misinterpret the emotion the other person is trying to convey. This is because emotional expressions are often subtle. A slight raising of the eyebrow or a quick smirk can be easily missed. Similarly, a sudden shift in tone is often hard to detect.

Although these gestures are subtle and easy to miss, they carry a wealth of information that can add context to the conversation. Detecting anger or sadness, for example, can help a human being respond appropriately to each interaction. Infusing this skill in machines can help create more useful and empathetic digital agents in the future. A machine that can detect emotions can generate responses that genuinely help users seeking assistance or information.

Machines are already detecting emotions in voice recordings and facial portraits. However, humans today, increasingly communicate using messaging applications. Their interactions through these applications are text-based. Popular social applications such as Whatsapp and Twitter encourage communication through short text messages. Without adequate context and additional information, machines can struggle to detect emotions in such conversations. For example, on reading "Why don't you ever text me!" you may either interpret it as a sad or angry emotion. The same ambiguity exists for machines. Lack of facial expressions and voice modulations make detecting emotions from text a challenging problem.

Moreover, as digital agents gain popularity in our society, it is essential that these agents are emotion-aware and respond accordingly. To solve this problem, we deployed a deep learning algorithm.

### Training wheels for AI

High-quality and high-volume data combined with appropriate labels is essential when solving a machine learning problem like this. The team decided to gather data in the form of Tweets sent out between 2012 and 2015. Of the hundreds of millions of tweets collected in this process, 300 were individually labelled by human judges.



Working with snippets of 3-turn conversations, we labelled each under one of four categories: *Happy*, *Sad*, *Angry* or *Others*. The human judges used textual cues, smileys, emoticons, and punctuation as clues to detect emotions throughout the set. For example, an exclamation mark after a word could often convey anger ("why!!"). Similarly, a smiley or emoticon such as ":)" could convey happiness. Often verbal cues can also indicate an emotion. For example, if a person responds to a text with "there, there," it could convey a need to comfort someone which may indicate a sadness or despair in the previous text.

Microsoft Edge

(<https://blogs.msdn.microsoft.com/msind/tag/microsoft-edge/>)

Convolutional Neural Net

(<https://blogs.msdn.microsoft.com/msind/tag/convolutional-neural-networks/>)

twitter

(<https://blogs.msdn.microsoft.com/msind/tag/twitter/>)

Artificial General Intelligen

(<https://blogs.msdn.microsoft.com/msind/tag/artificial-general-intelligence/>)

LVPEI

(<https://blogs.msdn.microsoft.com/msind/tag/lvpei/>)

Cloud Computing

(<https://blogs.msdn.microsoft.com/msind/tag/cloud-computing/>)

Recurrent Neural Network

(<https://blogs.msdn.microsoft.com/msind/tag/recurrent-neural-network/>)

CNNs

(<https://blogs.msdn.microsoft.com/msind/tag/cnns/>)

RUUH

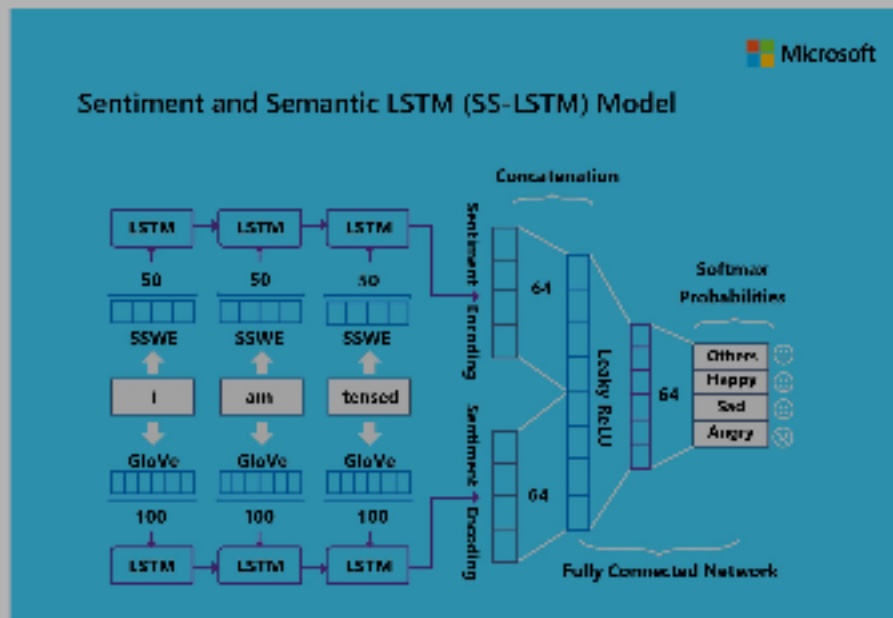
(<https://blogs.msdn.microsoft.com/msind/tag/ruuh/>)

### Recent Posts

- Making Ruuh, our AI-powered desi chatbot, see and respond to photos like humans (<https://blogs.msdn.microsoft.com/msind/2018/07/12/ruuh-ai-image-recognition/>) July 12, 2018
- Here's how Bing predicts possible outcomes of leading global sporting events (<https://blogs.msdn.microsoft.com/msind/2018/07/05/bing-predicts-sports-football-matches-ai/>) July 5, 2018
- Enabling accessibility in Microsoft Visio (<https://blogs.msdn.microsoft.com/msind/2018/05/17/creating-accessible-diagrams-inclusive->

Once the human judges pruned this relatively small dataset, a nearest neighbour based clustering algorithm was used to automatically sort the larger set into respective categories. The team finally obtained 456K utterances in the Others category, 28K for Happy, 34K for Sad, and 36K for Angry.

Massive data and appropriate labelling acted as the two training wheels for our Machine Learning solution for emotion detection. We called our solution 'Sentiment and Semantic Long Short-Term Memory (SS-LSTM) Model'.



As the name suggests, this Deep Learning model combines semantics and sentiment indicators to categorize text conversations based on the emotions they convey. To train the model, we used the Microsoft Cognitive Toolkit and divided the data into 'training' and 'validation', based on a ratio of 9:1 (9 sets for training and 1 set for validation). We discovered that the optimal batch size for training the model was 4000. With this batch size the learning rate of 0.005 gave us best results.

Finally, the team decided to test the model and compare its predictive power against other techniques. 2,226 3-tum conversations from 2016 were picked from Twitter to test the model.

#### The results

For each of the four emotion classes, our SS-LSTM model outperformed all other known state of art techniques for detecting emotions in textual conversations. The model performed better than individual LSTM-SSWE (Sentiment Specific Word Embedding) and LSTM-GloVe. SS-LSTM was also significantly better than Convolutional Neural Network (CNN) based approaches. Also, our Deep Learning based approach was found to be better than traditional Machine Learning techniques like Naive Bayes, Support Vector Machines and Gradient Based Decision Trees.

The key difference was the model's ability to detect emotions by combining semantics and sentiment indicators. In other words, the model simply 'understood' the conversation on a deeper level than other models.

#### Final thoughts

Conversations online are mostly in the form of short text messages. Without the benefit of voice modulation or facial expressions, detecting the emotion in conversations can be difficult. Despite the challenge, creating a digital agent that can detect emotions could be incredibly useful. Digital agents of the future may be a lot more capable if they can empathize with their users' feelings and respond appropriately. They could provide emotional support, contextual information, and even generate responses that suit the mood of the conversation.

By creating a model that combines semantics and sentiment indicators in short, text-based conversations, our team has helped take a significant leap forward in detecting emotions. The job is not done, but we're committed to keep chipping at it one bit at a time!

Tags AI (<https://blogs.msdn.microsoft.com/msind/tag/ai/>) artificial intelligence

(<https://blogs.msdn.microsoft.com/msind/tag/artificial-intelligence/>) CNN

(<https://blogs.msdn.microsoft.com/msind/tag/cnn/>) Convolutional Neural Network

(<https://blogs.msdn.microsoft.com/msind/tag/convolutional-neural-network/>) deep learning

(<https://blogs.msdn.microsoft.com/msind/tag/deep-learning/>) DL (<https://blogs.msdn.microsoft.com/msind/tag/dl/>) Emotions (<https://blogs.msdn.microsoft.com/msind/tag/emotions/>) Gradient

experiences-  
microsoft-visio/) May  
17, 2018

- Enabling industry 4.0 and smart factories with IoT and AI (<https://blogs.msdn.microsoft.com/msind/2018/04/27/10/smart-factories-ai-azure-iot-edge/>) April 27, 2018
- CodeTalk: Making IDEs more accessible for developers with visual impairment (<https://blogs.msdn.microsoft.com/msind/2018/04/10/code-talk-ide-accessibility-developers-visual-impairment/>) April 10, 2018

Based Decision Trees (<https://blogs.msdn.microsoft.com/msind/tag/gradient-based-decision-trees/>) IIT Kharagpur (<https://blogs.msdn.microsoft.com/msind/tag/iit-kharagpur/>) LSTM-GloVe (<https://blogs.msdn.microsoft.com/msind/tag/lstm-glove/>) LSTM-S-SWE (<https://blogs.msdn.microsoft.com/msind/tag/lstm-sswe/>) machine learning (<https://blogs.msdn.microsoft.com/msind/tag/machine-learning/>) Microsoft Cognitive Toolkit (<https://blogs.msdn.microsoft.com/msind/tag/microsoft-cognitive-toolkit/>) Naive Bayes (<https://blogs.msdn.microsoft.com/msind/tag/naive-bayes/>) Radhakrishnan Srikanth (<https://blogs.msdn.microsoft.com/msind/tag/radhakrishnan-srikanth/>) Sentiment and Semantic Long Short-Term Memory (<https://blogs.msdn.microsoft.com/msind/tag/sentiment-and-semantic-long-short-term-memory/>) Smileys (<https://blogs.msdn.microsoft.com/msind/tag/smileys/>) SS-LSTM (<https://blogs.msdn.microsoft.com/msind/tag/ss-lstm/>) Support Vector Machines (<https://blogs.msdn.microsoft.com/msind/tag/support-vector-machines/>) twitter (<https://blogs.msdn.microsoft.com/msind/tag/twitter/>) Umang Gupta (<https://blogs.msdn.microsoft.com/msind/tag/umang-gupta/>) WhatsApp (<https://blogs.msdn.microsoft.com/msind/tag/whatsapp/>)

## Comments (2)

---

Name \*

Email \*

Website

Post Comment



*Stacie*

November 2, 2017 at 10:00 pm (<https://blogs.msdn.microsoft.com/msind/2017/10/06/humanizing-artificial-intelligence-ai-with-deep-learning/#comment-605>)

Hello, is there someone I can email for more information on this subject and future research directions?

Thank you!

Reply (<https://blogs.msdn.microsoft.com/msind/2017/10/06/humanizing-artificial-intelligence-ai-with-deep-learning/?replytocomment=605#respond>)



*Shrishail Pathani*

November 14, 2017 at 5:07 am (<https://blogs.msdn.microsoft.com/msind/2017/10/06/humanizing-artificial-intelligence-ai-with-deep-learning/#comment-2057>)

Really interesting way to apply emotion-detection to text based conversation. We tried built pilot implementation of Chatbot solution, forget about detecting the emotions, giving it somewhat closer to human conversation was hard. We had to build at least a thousand utterances to train the model and kept on fixing unresolved conversations via LUIS model to make it work. Love to hear about your progress on this topic.

Reply (<https://blogs.msdn.microsoft.com/msind/2017/10/06/humanizing-artificial-intelligence-ai-with-deep-learning/?replytocomment=2057#respond>)

