WFH BIOMETRIC FACIAL EMOTION RECOGNITION

While we have a lot of quantitative metadata floating around today, across all industries, there still remains a disconnect in most AI applications between that data and human emotions. But in the broader scope of technology, there is an ever increasing ability to process large amounts of data, i.e. what the agent (and/or customer) is saying and what emotion they are expressing nonverbally. If we develop a system that translates this agent data (emotional state), we can see benefits across many areas in the contact centers, including improvements to the workplace environment, coaching/training, product development, and other critical areas like fraud detection and customer security.

That being true, Capital One has a need to remain competitive in the crowded digital market, and human emotion has a significant role to play as a unique business advantage. We have a need for data; a large portion of current economic and scientific progress depends on the availability of data about individuals' behaviors - as this is defined, our contact center has a lot of data that goes unanalyzed. We also have a future need of a "valuation/productivity score" applied to each agent as long term we think about the majority of phone agents with bots.

Background

Facial emotion recognition - separate from the standard facial recognition - came to light upon the realization that people are not as binary as machines and the continual interplay of human emotions is the foundation for informational processing.

Emotion AI is based on the simple principle of input -> output. This type of artificial intelligence is designed (parts both proven and theoretical) to recognize, identify, process and translate human affects and the system should be able to provide an appropriate response - whether that be automated alerts and reporting or real time feedback. In both B&M environments and WFH, Capital One offers hardware and software that with advancements in technology, could be utilized in this implementation, primarily specific webcams and headsets and specific desktop security programs.

Current AI programs exist that measure aspects of emotion such as tone of voice, pitch, and cameras in our virtual environment have the capability to measure other aspects

such as posture, gestures, micro expressions which combined can even go as far as to translate the rhythm or force of the keystrokes and temperature of the hands into quantifiable data that can be tracked. With those advancements, cameras and sensors can take that input data and take it though deep learning algorithms (that already exist) to generate a valid response. Re scope and scalability, this concept is limitless, as virtually even human behavior is driven by human emotion.

PROPOSAL: Use biometric Facial Emotion Recognition in a process that maps out facial expressions of our WFH Call Center agents to identify emotions, deriving a quantitative productivity and valuation score

WFH Biometrics facial recognition - emotion AI - can guide our agents to better performance. The contact center is an excellent environment to leverage emotion AI, combined with audio mining (speech analytics). Agents take calls continuously, so monitoring our WFH agents' emotions can not only help them handle the call better, but also guarantee a more effective customer experience with potential for reduced costs and increased income.

Proposal to model our agent productivity for our Bank Contact centers based on biometric feedback

Today, is it hard to accurately measure the productivity of a contact center agent as measures commonly used for evaluation are primarily subjective. For example, coaching feedback is holistically subjective, and an agents 'tone' on a call can technically be either a matter of fact, or more commonly, something that is perceived by the listener. These subjective factors create bias and can lead to performance evaluations that are inconsistent, versus being based on their real or actual performance. An agent's personal style and individuality can give a lot of room for interpretation.

The goal of assessing a productivity score is segmenting an agent into a binary classification, either productive or nonproductive

This proposed productivity score, an objective performance review, is an extension of our contact center standards that an agent follows, supported by their product and technical knowledge in responding to customers inquiries, a common example is scripting our calls based on predefined scenarios. This score would be calculated from facial emotion (biometric expression) data from webcams, combined with audio mining technology from headsets as a way to use machine learning and affective computing for predictive analysis of the agent.

When we think about the benefit of such a productivity score, it's important to note that contact centers are dynamic and inundated with customer contact over various communication channels, so our organization's bottom line performance is tied directly to performance of the agents handling those contacts. If we automate the detection of facial emotions and tie those back to predefined criteria the model follows, we can reduce effort and time in a manual evaluation.

There have been recent developments in text-based information (call transcript keyword filters), and speech recognition (audio mining). Evaluation productivity of an agent from the emotional signals, coupled with speech signals, is a vast improvement from our current text-based approach used in Converse. The data extracted from biometric innsights carries much more information when compared to transcribed text.

Intelligent Voice and Audio Recognition model
I recommend a modeling approach using Bayesian Information Criterion (BIC vs AIC) , Hierarchical Agglomerative Clustering (HAC), and Viterbi computing.

The concept of this model is that the communication process is represented by four elements (emoter, distal cues, proximal percepts, and observer) and three phases (externalization, driven by external models and internal changes; transmission cue utilization, driven by inference rules; and schematic recognition). Without going into technical detail of the underlying algorithm, if we apply this model to emotion AI, the internal state of the agent is encoded by what are known as "distal vocal or facial cues" and the listener ('us') perceives the utterances and extracts several cues (e.g. voice

quality ratings or pre-mapped emotional patterns). It is these perceptions that are used to gather an internal state of the agent based upon a set of rules.
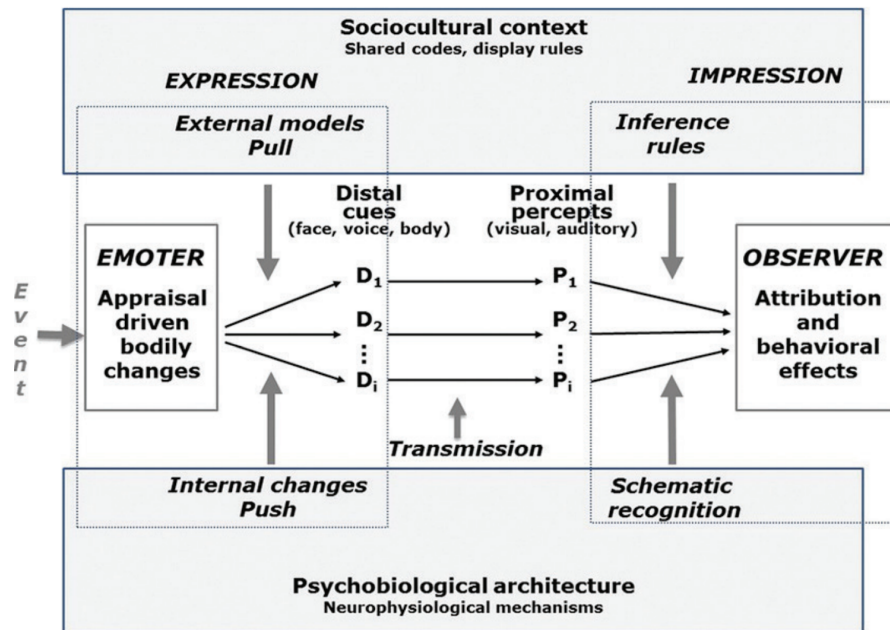
Additional Applications: Understanding emotional intelligence is a key skill
- Training
  - Our e-learning can utilize emotion AI to detect when an agent is having difficulty or needs guidance
- Coaching/Working environment
  - By understanding the objective emotions of their agents, coaches can better foster a positive connection (rapport), regulate any personal feelings in potentially hostile communication, provide real time call feedback and make sure every agent is treated appropriately.
- Customer security and fraud
  - Using emotion AI, our fraud teams can build out machine learning to better predict fraudsters, based on signals from the agent (fear, suspicion, threat)
- Adapt to a post-COVID-19 world

- - - It's expected that we'll have fewer touch screens and more voice interfaces and machine vision interfaces, so implementing facial emotion recognition is next step
  - Product Development
    - Analyze agent emotions to display personalized messages in smart environments like in E4A or Pinpoint, such as reminders of critical procedures such as confusion

---

Conclusion

With rapid advancement in technology, big data, AI and cloud computing, the understanding and predicting of human emotions can be realized. Speech, facial, and movement analysis algorithms have grown powerful enough to be able to analyze all features of interactions in a granular detail, creating a number of benefits.

Facial expression recognition could be one of the most powerful, natural and immediate means for human beings to communicate their emotions and intentions. Using biometric facial emotion recognition on our WFH agents intersects AI/ML, big data, human computer interaction and psychology that can inform stakeholders of critical performance indicators critical to future contact center growth.

While there are limitations and ethical considerations, it is expected that these will be addressed as the technology continues to develop. The emotion detection and recognition market is highly investable and Capital one needs to take heed of its capabilities, and place a focus on the potentially lucrative drivers of emotional intelligence.