

A IMPORTÂNCIA DO APRENDIZADO DE MÁQUINA NOS ESTUDOS DOS AFETOS

THE IMPORTANCE OF MACHINE LEARNING IN AFFECT STUDIES

Saulo Almeida[†], & Paulo Caetano da Silva¹

¹Universidade Salvador, Brasil.

Resumo: Aprendizado de máquina é um campo da Inteligência Artificial no qual um sistema pode modificar um comportamento, de forma autônoma, baseando-se na sua própria experiência, com o uso de regras lógicas que modificam o desempenho de uma tarefa ou tomam a decisão mais apropriada para o contexto. Um dos seus principais desafios é pensar como uma máquina pode aprender a lidar com afetos ou seus correlatos na interação com seres humanos ou outros agentes. Para isso, este trabalho busca compreender o que se tem produzido em uma associação entre as áreas dos afetos e do aprendizado de máquina, através de uma revisão sistemática de literatura. Após as filtragens aplicadas, discute-se 22 artigos que tratam dos temas de pesquisa. Em geral, os trabalhos lidam diretamente com a afetividade humana ou esta aparece como elemento ligado aos conceitos gerais, como é no caso dos estudos acerca do engajamento. Os trabalhos discutidos se destacam pelo uso de aprendizado de máquina na leitura e análise de dados retirados de resultados de exames diagnósticos ou de redes léxicas, associados a equipamentos de análise da atividade cerebral, destacando o eletroencefalograma (EEG), ou mesmo, o uso de robôs, *chatbots* ou jogos eletrônicos. *Palavras-Chave:* Aprendizado de máquina, Afetos, Inteligência artificial, Avaliação psicológica, Psicologia.

Abstract: Machine learning is a field of Artificial Intelligence in which a system can modify a behavior, autonomously, based on its own experience, using logical rules that modify the performance of a task or make the most appropriate decision. for the context. One of its main challenges is to think about how a machine can learn to deal with affects or their correlates in the interaction with human beings or other agents. For this, this work seeks to understand what has been produced in an association between the areas of affection and machine learning, through a systematic literature review. After the filters applied, 22 articles dealing with the research topics are discussed. In general, the works deal directly with human affectivity or it appears as an element linked to general concepts, as is the case of studies on engagement. The works discussed stand out for the use of machine learning in the reading and analysis of data taken from the results of diagnostic tests or lexical networks, associated with equipment for analyzing brain activity, highlighting the electroencephalogram (EEG), or even the use of robots, *chatbots* or electronic games.

Keywords: Machine learning, Affections, Artificial intelligence, Psychological assessment, Psychology.

O desenvolvimento tecnológico, nos últimos anos, revelou a possibilidade de uma humanidade construindo relações sociais, cognitivas e afetivas com ou através do uso de máquinas auxiliaadoras

[†]Autor de Correspondência: Saulo Almeida (professorsauloalmeida@gmail.com)

Submetido: 23 de março de 2023

Aceite: 05 de dezembro de 2023

dos trabalhos humanos. Aquilo que parecia algo distante, presente apenas nos filmes de ficção científica, começa a ficar cada vez mais próximo.

No entanto, alguns limites são impostos neste processo de desenvolvimento de “máquinas pensantes”. Enquanto o conhecimento das máquinas acontece por meio de um raciocínio lógico e analítico, o conhecimento/aprendizado humano ocorre de forma intuitiva, com o uso de emoções e afetos. Nesse contexto, os estudos acerca da tecnologia e da cognição humana devem levar em consideração a questão de como as máquinas podem se assemelhar às condições humanas. Como sustentar uma relação socioafetiva com o uso de instrumentos e ou máquinas?

Para Picard (1997) isto é possível ao analisar um novo campo de estudos em computação: a computação afetiva. Em uma computação mais afetiva, os estudos dos afetos e emoções nas máquinas são o cerne de toda e qualquer discussão, seja através da inserção das emoções humanas na máquina, seja no reconhecimento das emoções humanas, podendo ser, portanto, um avanço do aprendizado de máquina. Neste sentido, este processo pode ser mais bem analisado ao se destacar que o aprendizado de máquina (ou seu termo em inglês *machine learning*) está cada vez mais próximo do nosso cotidiano.

Pedro Domingos no início do prólogo do seu livro “*O Algoritmo Mestre: Como a Busca pelo Algoritmo de Machine Learning Definitivo Recriará Nosso Mundo*” afirma que talvez o aprendizado de máquina esteja em todos os locais ao nosso redor (Domingos, 2019). Pode parecer exagero, mas desde quando se busca algum item em um mecanismo de busca na internet, até mesmo ao intercambiar e-mails, o uso do aprendizado de máquina está presente.

Aprendizado de Máquina é uma técnica utilizada em sistemas de informação o qual pode modificar um comportamento, de forma autônoma, baseando-se na sua própria experiência, com o uso de regras lógicas que modificam o desempenho de uma tarefa ou tomam a decisão mais apropriada para o contexto, ou seja, são programas capazes de aprender por si sós, utilizando-se de um conjunto de dados. É, deste modo, uma área de pesquisa multidisciplinar que engloba inteligência artificial, psicologia, neurobiologia, entre outros (Cerri & Carvalho, 2017). Para isto, o aprendizado de máquina se utiliza de programas de computador que aprendem e fazem previsões sobre dados a partir de abordagens de aprendizagem, como supervisionada, não supervisionada, semi-supervisionada e por reforço.

O aprendizado de máquina supervisionado é um método cuja forma de aprender é a partir de experiências predefinidas, entradas de dados rotulados, com suas respectivas classes (saídas), para aprender padrões e classificar novas entradas. No aprendizado não supervisionado não existem experiências predefinidas (dados de entrada não rotulados), aprendendo e detectando agrupamentos implícitos usando-os para a categorização. Obviamente, o semi-supervisionado abrange ambos os métodos: supervisionado e não supervisionado, utilizando tanto o conjunto de dados rotulados como também o conjunto de dados não rotulados, é o aprendizado por reforço, onde o aprendizado da máquina pode ser moldado por estímulos e respostas.

Diante destas possibilidades, entende-se como desafio a implicação de uma lógica cognitiva da capacidade intelectual humana aplicável ao contexto das máquinas. A partir desta constatação, algumas perguntas são feitas: Como lidar com questões subjetivas? Pode a máquina aprender a lidar com afetos ou seus correlatos, como sentimentos e emoções? Além do campo da aplicação do modo de ser humano às máquinas, como estas também podem auxiliar o processo de significar o mundo?

Portanto, se o comportamento, os afetos e o aprendizado são características primariamente humanas, há de se imaginar uma possibilidade de contato entre aprendizado de máquina e estes campos humanos, pois ambas as áreas lidam com o comportamento, só que a primeira mais diretamente associada a máquinas e não, especificamente, ao comportamento humano. Neste sentido, o que se questiona neste trabalho é como lidar com aquilo que escapa a um olhar experimental: o campo dos afetos. Assim sendo, destaca-se a necessidade de investigar sobre o campo dos afetos

quando se trata de aprendizado de máquina, a partir da questão de pesquisa: Como o campo dos afetos tem sido trabalhado no âmbito dos estudos em aprendizado de máquina?

MÉTODO

Na busca de compreensão do que se tem produzido em uma associação entre as áreas da psicologia, no que tange aos afetos, e do Aprendizado de Máquina, escolheu-se a revisão sistemática de literatura como metodologia desta pesquisa. Para realizar a revisão sistemática da literatura adotou-se uma metodologia muito utilizada no domínio da Engenharia de Software (De-la-Torre-Ugarte-Guanilo et al., 2009; Dermeval et al., 2020). Essa metodologia consiste na definição das questões de pesquisa sobre aquilo que se deseja investigar, para em seguida se definir os protocolos de pesquisas, i.e. *strings* de busca, fontes, critérios para inclusão e exclusão de trabalhos relacionados, critérios de qualidade para seleção dos artigos e análise dos trabalhos para entendê-los e encontrar respostas às questões de pesquisa. A seguir é mostrado como foi aplicada esta metodologia neste trabalho.

Para escolha das *strings* de pesquisa, foi necessário executar pesquisas de teste com as *strings* de pesquisa candidatas, usando os seguintes repositórios: IEEE Xplore, Science Direct, ACM Digital Library, PubMed, Biblioteca Virtual de Saúde (BVS), Association for Information Systems (AIS) e Springer. Esses repositórios foram escolhidos porque eles estão entre os mais utilizados em mapeamentos e revisões sistemáticas na área da computação e da saúde.

Como resultado do procedimento descrito, foi definida a *string* de busca: Machine Learning AND Psychology AND Affection. Os seguintes critérios de inclusão foram aplicados aos estudos obtidos a partir da *string* de pesquisa: O trabalho foi desenvolvido nos últimos 5 anos (trabalhos mais recentes); O trabalho apresenta Afeto, ou sinônimos, no seu título ou palavras-chaves; O artigo trata de aplicações do aprendizado de máquina na área da psicologia; Serão aceitos estudos primários e secundários; Se vários artigos de periódicos relatam o mesmo estudo, o mais recente é incluído.

Como critério de exclusão, o seguinte foi aplicado: O trabalho apresenta a discussão em apenas uma das áreas, sem levantar uma integração entre ambas. Além dos critérios gerais de inclusão e exclusão, os seguintes critérios de qualidade, definidos pelos pesquisadores, foram aplicados no processo de seleção, mantendo-se apenas no mapeamento sistemático os estudos para os quais todas as respostas foram afirmativas: Os objetivos da pesquisa estão claramente declarados? Os resultados ou conclusões estão claramente explicados? Os resultados ou conclusões obtidas estão relacionados com as questões de pesquisa? O plano de estudo é adequado para responder às questões de pesquisa?

RESULTADOS

O conjunto base de estudos foi obtido a partir da aplicação das *strings* de pesquisa nos bancos de dados científicos. Para um melhor entendimento dos resultados de busca, inicialmente separou-se os artigos por base de dados. Foram aplicados, então, os critérios de filtragem presentes na Figura 1.

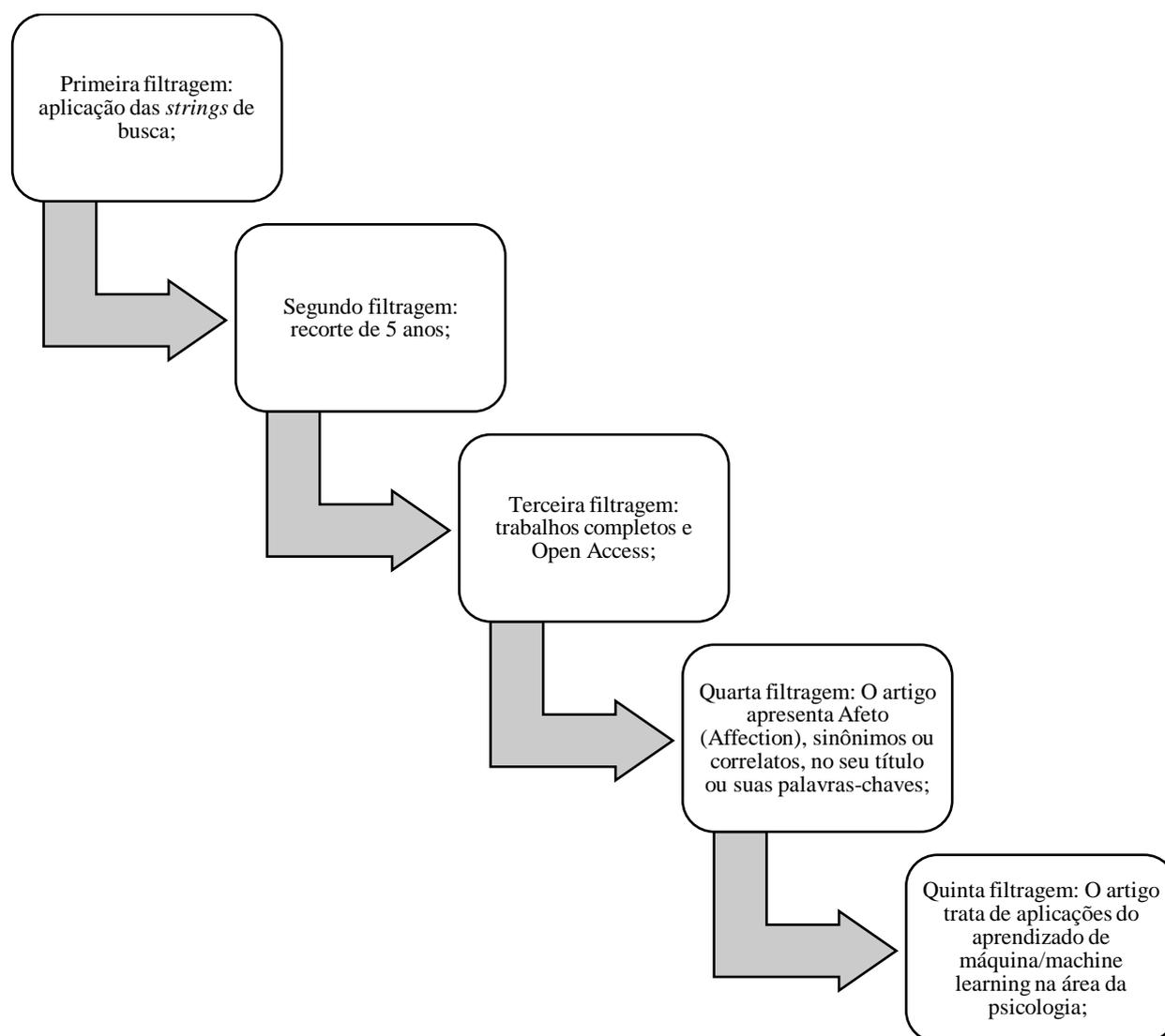


Figura 1. Critérios de filtragem aplicados na pesquisa

Observa-se que alguns artigos encontrados na pesquisa geral com as *strings* discutiam acerca de inteligência artificial, ou mesmo o uso de robôs ou *chatbots* (programa de computador que tenta simular um ser humano na conversação com as pessoas, muitas vezes utilizando inteligência artificial para obter uma experiência mais próxima do real), mas não tinha o aprendizado de máquina como aspecto central da sua discussão.

Na base Science Direct foram encontrados 831 trabalhos após a primeira filtragem. Após as filtrações subsequentes, restringiu-se para 6 trabalhos. Na base ACM Digital Library foram encontrados 572 trabalhos a partir das *strings* de busca, sendo que ao aplicar os filtros de recorte dos últimos 5 anos reduziram-se para 273. Na IEEE Xplore encontram-se apenas 4 artigos a partir do uso das *strings* deste trabalho, sendo 3 conferências e um artigo. No PubMed apenas um trabalho foi identificado depois das filtrações de recorte temporal e trabalhos abertos, porém quando aplicado como tema central o Afeto, não pode ser selecionado nenhum trabalho. Na base de dados a Biblioteca Virtual de Saúde (BVS), aplicadas as três primeiras filtrações restou apenas 1 artigo, que se manteve até a última filtragem. Na AIS encontram-se 2514 trabalhos com as *strings* de busca, sendo que 1195 nos últimos 5 anos e 146 completos.

No total tem-se 9184 trabalhos encontrados, porém com as filtrações aplicadas, restaram 22 artigos que tratam dos temas de pesquisa, completos e gratuitos para *download*. O Quadro 1 mostra um resumo dessas buscas após a aplicação dos critérios.

Quadro 1. Trabalhos encontrados por filtragem – strings

Bases de dados	<i>Machine Learning AND Psychology AND Affection</i>				
	Triagens				
	1ª filtragem	2ª filtragem	3ª filtragem	4ª filtragem	5ª filtragem
<i>Science Direct</i>	831	220	42	13	6
<i>ACM Digital Library</i>	572	272	28	9	3
<i>IEEE Explore</i>	4	3	0	-	-
<i>PubMed</i>	1	1	1	0	-
<i>BVS</i>	2	2	1	1	1
<i>AIS</i>	2514	1195	146	10	1
<i>Springer Link</i>	5262	2085	75	57	11
Total	9184	3778	294	91	22

DISCUSSÃO

Esta seção apresenta a análise feita sobre os artigos selecionados em relação a cada questão de pesquisa, de forma a explicitar os dados mais relevantes para atingir os objetivos deste trabalho – identificar os objetivos dos estudos, analisar as principais dificuldades de lacuna, e identificar os afetos mais estudados. Apresenta-se a descrição dos artigos conforme seus objetivos de pesquisa, destacando a relação entre a pesquisa, ou revisão, e o campo dos afetos.

O artigo de Lim et al. (2022) discute o comportamento emocional do consumidor a partir de uma revisão de literatura através da combinação de análises bibliográficas e temáticas em 861 artigos que tratam de Customer Engagement (CE ou Engajamento do Cliente) publicados em 377 periódicos indexados na Scopus entre 2006 e 2020. Engajamento do cliente é o conceito principal utilizado para identificar as diferentes perspectivas do comportamento desejado do cliente, sendo que os afetos envolvidos são apresentados na construção do perfil, preferência e comportamento dos clientes. Este perfil é construído através do uso de aprendizagem de máquina, destarte, a afetividade é traçada a partir deste perfil, principalmente ao compreender as preferências/gostos dos clientes. Neste caminho, Forgas-Coll et al. (2021), com o uso de teorias psicológicas, como a Teoria da Ação Racional e o Modelo de aceitação da Tecnologia, propuseram testar um modelo de aceitação tecnológica para o caso de robôs sociais (robô autônomo que interage com humanos ou outros agentes autônomos, a partir de papéis definidos) após uma experiência real de entretenimento. Para isso, usaram o TIAGo, que é um robô semi-humanóide, com expressões faciais e coordenação com a conversa. Os resultados demonstraram que a maior influência do robô está relacionada a respostas socioemocionais, seguido por motivos hedonistas dos sujeitos usuários, porém há também fatores cognitivos como utilidade percebida e facilidade de uso percebida.

Ainda sobre a relação entre o uso de robôs e o eliciamento de emoções, o trabalho de Sutoyo et al. (2019) discute uma estrutura de *chatbot* emocionalmente realista para aprimorar sua credibilidade. Com Linguagem de Marcação de Inteligência Artificial (AIML - linguagem baseada em XML desenvolvida para criar diálogos semelhante a linguagem natural por meio de softwares, simulando assim inteligência humana), os autores partem do pressuposto de que a expressão das emoções, por um objeto (humano, animal, desenho animado, *chatbot*, humano virtual etc.) aumentará a credibilidade do objeto. Criaram, então, um *chatbot* onde o usuário insere um texto via teclado e as emoções também proporcionam conversas com cores de fontes coloridas nas interações. Eles revelam nos resultados que há uma melhora significativa na credibilidade do sistema aprimorado (com modelos de emoções) em comparação com o *chatbot* normal (sem modelos de emoções), ou seja, há uma melhora estatisticamente significativa na credibilidade do *chatbot* no sistema que possui variáveis de emoções induzidas em comparação ao sistema sem emoções.

Outro artigo que também apresenta um *chatbot* é o de Zhou et al. (2020). Os autores descrevem o desenvolvimento do Microsoft XiaoIce, projetado para ser um companheiro emocional capaz de satisfazer a necessidade humana de comunicação, afeto e pensamento social. Para isso, o chatbot precisa identificar as emoções a partir de conversa, as mudanças ocorridas neste diálogo ao longo do tempo e, com isso, prever as necessidades emocionais do usuário. Este caminho requer formação de perfil, detecção de emoção, reconhecimento de sentimento e rastreamento dinâmico do humor. Nisto, o sistema gera um vetor de empatia, que consiste em uma lista de valores-chaves, gerados usando classificadores de aprendizado de máquina, que representam o perfil, as intenções, emoções e opiniões do usuário.

A transmissão de emoções é também analisada através do estudo de como os humanos transmitem emoções através do toque (Andreasson et al., 2018). Com o uso do robô NAO (Figura 2), os sujeitos são instruídos a imaginar que tinham determinadas emoções e que o robô deveria entender como a pessoa se sente ao tocá-lo.



Figura 2. Robô NAO. Fonte: RobotLAB - <https://www.robotlab.com/loja/robo-nao>.

Na Figura 3 é possível visualizar a classificação de Support Vector Machine (SVM; conjunto de métodos de aprendizado supervisionado que analisam os dados e reconhecem padrões, usado para classificação e análise de regressão) da valência emocional. Ao usar o Matlab (software interativo de alta performance voltado para o cálculo numérico) para a classificação SVM bidimensional, os valores médios foram analisados a partir do número de diferentes locais tocados e a duração do toque, para classificar as emoções. Os resultados mostraram que as emoções são transmitidas diferencialmente ao longo de várias dimensões – intensidade, duração, localização e tipo. Os principais resultados encontrados são as diferenças de gênero: as mulheres transmitem emoções através do toque ao robô por durações mais longas do que às dos homens e com uma distribuição maior de locais. Quanto às emoções, a *tristeza* é a que tem o toque mais demorado e o *amor* é a emoção que se transmite com mais distribuição de toques.

Os trabalhos de Papanastasiou et al. (2020), Kocon e Maziarz (2021), Chae e Lee (2022), Sutoyo et al. (2019), Frewen et al. (2020) exploram a questão das emoções ao tratar de pesquisas acerca da Interface Computador-cérebro (Brain Computer Interface – BCI), bem como práticas de intervenção para treinamento e reabilitação de alunos com transtornos do neurodesenvolvimento. Esses autores revelam avanços na tecnologia de eletroencefalograma em técnicas computacionais e aprendizado de máquina, através de sistemas de reconhecimento baseados em emoções que melhoram a comunicação entre usuários e máquinas, e experiências de usuário que permitem identificar níveis de meditação, engajamento, frustração, excitação e estresse. Como exemplo, eles citam uma técnica de eliciar emoções baseada nos neurônios-espelhos (Mirror Neuron System), que é um sistema de neurônios especializados que “espelha” as ações e o comportamento dos outros, impactando funções neuro cognitivas ou distúrbios neuropsiquiátricos. Esta técnica permite o monitoramento avançado dos

status de usuários em relação as seis emoções básicas: felicidade, surpresa, raiva, medo, nojo e tristeza (Petranonakis e Hadjileontiadis, 2010). Neste modelo, estas pesquisas revelam que intervenções que utilizam a interação homem-computador, principalmente o uso de jogos e neuro feedback, mostram resultados importantes no desenvolvimento de níveis de criatividade, diminuição da ansiedade, e entendimento sobre as emoções eliciadas.

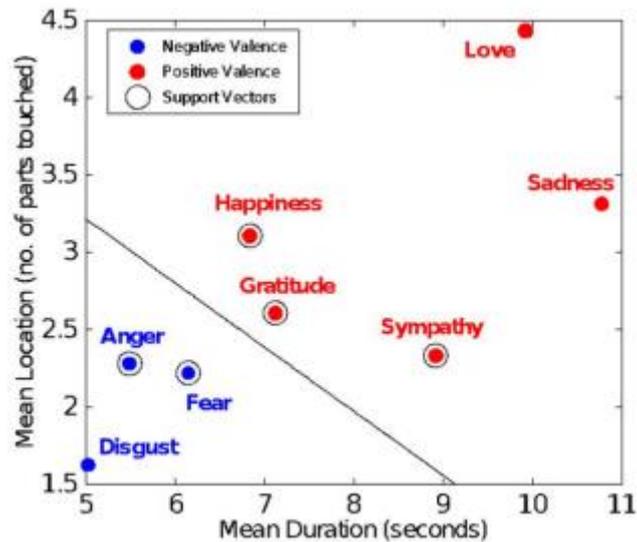


Figura 3. Classificação de Support Vector Machine. Fonte: Andreasson et al. (2018).

No reconhecimento de emoções a partir da interface Humano-computador, o uso da fala é também um tópico amplamente pesquisado, como cita Krishnan et al. (2021). O artigo desses autores investiga uma tentativa de reconhecimento de sete estados emocionais a partir de sinais de fala (triste, zangado, desgostoso, feliz, surpreso, agradável e neutro). Eles fazem isso dividindo os componentes de frequência no sinal de fala em três grupos: alta frequência, média e baixa frequência. Com estas diferenças de frequência, pode-se reconhecer emoções existentes em um sinal de fala.

Esta discussão é levada também para o campo do toque interpessoal afetivo. Ocklenburg et al. (2022), com base nos avanços em etologia computacional em modelos animais, demonstram que ao utilizar métodos de aprendizado de máquina para rastrear e quantificar comportamentos naturais complexos, e analisar o toque social, como o abraço, tem-se a necessidade de uma integração de dados de etologia computacional humana com dados coletados usando dispositivos portáteis de medição neurocientífica, como o eletroencefalograma.

Quanto a diagnósticos, Frewen et al. (2020) reforça a importância do uso do aprendizado de máquina em dados de imagem estrutural do cérebro, capaz de prever não apenas doenças, mas também síndromes. Um exemplo é o uso do aprendizado de máquina para a leitura e análise de dados no diagnóstico de doença de Parkinson e síndromes do membro alienígena (doença rara que gera a sensação de perda de controle dos membros e realização de movimentos involuntários.). O ensaio revisa a neuroimagem funcional através de pesquisas que tem o objetivo de investigar o processamento autorreferencial (SRP), ou seja, como respondemos a estímulos que nos referenciam. Neste caminho de diagnósticos, Antonucci et al. (2021) preocupam-se com o diagnóstico da psicose. Defendendo a ideia de que as relações parentais e o desenvolvimento de um apego inseguro durante o desenvolvimento humano podem ser fatores ambientais provocadores de uma vulnerabilidade para psicose, os autores buscam criar um algoritmo de aprendizado de máquina baseado na qualidade percebida dos recursos relacionados ao estilo parental e apego, a fim de rastrear a condições de risco e/ou estágios iniciais de psicose. Com o uso de instrumentos psicométricos de avaliação da cognição social e das emoções dos sujeitos, o sistema de aprendizado de máquina foi alimentado, e ajudou na atribuição de causalidade de psicose associada a percepções negativas da infância e apego inseguro,

mesmo que com restrições de uma amostra homogênea (caucasianos). A estratégia geral de aprendizado de máquina foi realizada por meio do software NeuroMiner, versão 1.

Hu (2021) examina a relação entre as adversidades da infância e o envelhecimento saudável na população chinesa. Usando os dados da China Health and Retirement Logitudinal Survey, uma pesquisa nacional que coletou a história de vida e informações relacionadas ao envelhecimento de 9.248 idosos com mais de 60 anos em 2014 e 2015, através de agrupamento k-means, uma técnica de aprendizado de máquina não supervisionado. Com os indicadores de saúde normalizados em um intervalo entre 0 e 1, investigou-se as associações entre as adversidades da infância e a variável categórica envelhecimento saudável. O estudo mostra que as questões de gênero atravessam a possibilidade de um envelhecimento saudável, e isso não se limita a um estágio específico da vida, mas abrange todo o curso da vida, e as adversidades da infância influenciam o envelhecimento saudável, porém pode ser atenuada à medida que as pessoas atingem uma idade mais avançada.

Chae e Lee (2022), no campo das emoções positivas e negativas, examinaram como a pandemia do COVID-19 se associa à emoção positiva, emoção negativa e atitude em relação à interação social dos usuários da rede *Twitch* (serviço de streaming ligado a corporação Amazon, que se concentra na transmissão de competições de videogame), usando o processamento de linguagem natural (NLP) como método, através de dois conjuntos comparáveis de dados de texto: um após o surto da pandemia e outro antes disso. Eles encontraram emoções mais negativas, como raiva e ansiedade, como as mais significativas após o surto da pandemia. Para isso, utilizam análise baseada em dicionário através da técnica Contagem de Palavras (LIWC; recurso linguístico que permite calcular o grau de uso de diferentes categorias de palavras), em aprendizado de máquina supervisionado, para prever a emoção dos usuários.

Em termos de emoções propagadas nas redes, Kocon e Maziarz (2021) adaptaram um método para a propagação automática de anotação de emoção em redes de palavras. Os autores investigaram a possibilidade de aplicar o conhecimento da ciência neural para processamento de linguagem natural, ligando uma rede léxico-semântica com uma rede de áreas cerebrais interconectadas (HBC). Através de dois experimentos, sendo o primeiro com o objetivo de classificar palavras polonesas, extraídas do *WordNet* (banco de dados lexical de relações semânticas entre palavras em mais de 200 idiomas) e mapeadas através de sinônimos emocionalmente significativos ou neutros. Houve também o mapeamento cerebral e construção de vetores das probabilidades de ativação de regiões cerebrais, com o intuito de prever a polaridade emotiva da palavra. O segundo experimento foi um experimento em larga escala sobre a propagação de emoções em um ambiente bilíngue (Princeton WordNet e Polish WordNet), via relações semânticas interlinguais: sinônimos, hipônimos e hiperônimos). Notou-se que uma rede neural profunda treinada teve um desempenho muito melhor do que uma de base aleatória obtida através do embaralhamento de vetores cerebrais entre o conjunto de dados. Modelos de aprendizado de máquina baseados nos vetores HBC tiveram respostas positivas com linhas de base aleatórias.

Williams et al. (2019), no campo destas análises de linguagem, investigam diferentes esquemas de classificação da linguagem emocional. A partir de um corpus de 500 documentos de texto do Twitter, carregados de emoção, eles comparam seis esquemas: as seis emoções básicas de Ekman, a roda da emoção de Plutchik, a teoria do afeto de Watson e Tellegen, a Emotion Annotation Representation Language (EARL), WordNet–Affect (uma extensão dos Domínios WordNet, incluindo um subconjunto de synsets adequados para representar conceitos afetivos correlacionados com palavras afetivas), e o texto livre. Tratando-se de uma computação afetiva, destaca-se aqui a EARL, que é uma linguagem formal para representar emoções em contextos tecnológicos, projetado para anotação de corpus e reconhecimento de emoções, e organização das emoções como positivas e negativas, apuradas com base na intensidade e na atitude. A Figura 4 ilustra 48 emoções que podem ser mapeadas com a EARL.

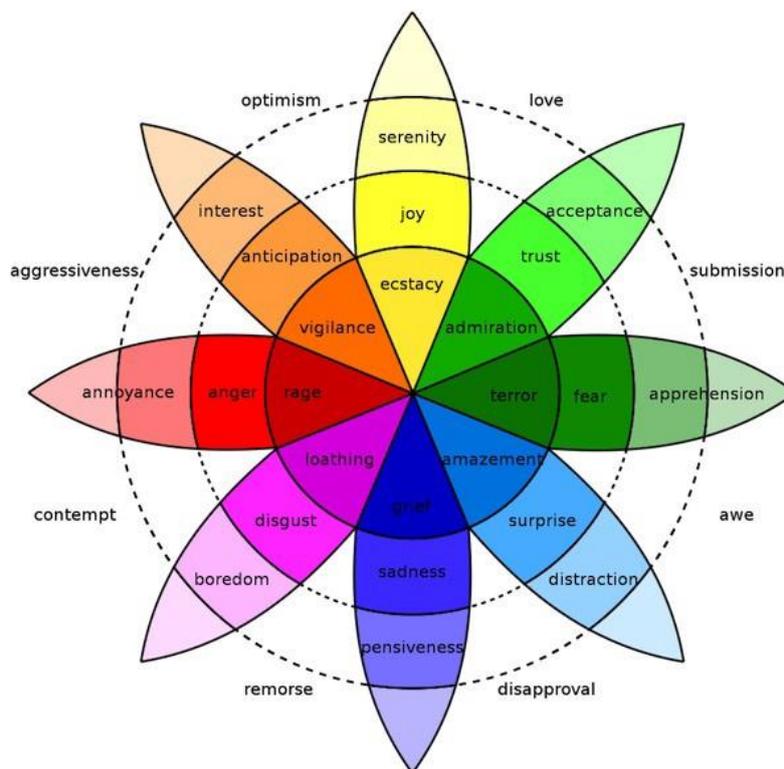


Figura 4. 48 emoções que alguém pode sentir segundo EARL. Fonte: Williams et al. (2019).

Nota-se, que o campo dos afetos/emoções na área do aprendizado de máquina é amplo e com várias possibilidades. No entanto, uma questão é importante ser destacada quando se trata de interação máquinas e seres humanos: a questão ética. Sætra (2020) amplia esta discussão ao sustentar a ideia de que os robôs são parasitas sociais, pois aprendem na interação social, exploram e se alimentam de comportamentos humanos. A preocupação é como lidar com a atenção conjunta, as intenções compartilhadas, o significado dialógico e a existência no triângulo referencial (criança, adultos e máquina), questionando-se até a autenticidade das relações humanas se comparadas com a relação homem-máquina. Se os robôs forem capazes de relacionamentos autênticos, pode-se inferir que os seres humanos são pouco mais que os robôs e ferramentas que os cercam, o que nos leva a pensar na construção de vínculos afetivos capazes de enraizar ou suplantar preconceitos e discriminações.

Hutson et al. (2018) discutem a projeção de sistemas técnicos que sejam resistentes a preconceitos e discriminações. Através de pesquisa em plataformas íntimas (plataformas de namoro e conexão online), eles questionam os recursos de design das plataformas que permitem a triagem com base em características, fornecendo aos usuários uma forma de controle da seleção de parceiros em potencial, naturalizando preferências discriminatórias.

A proposta de Ekman e Friesen é de que as plataformas façam uso de recursos de classificação padrão sem categorias sobrecarregadas de preconceitos e estereótipos, muitos ligados a questões de raça, sexo, idade e habilidades. Isso em meio ao desafio de prever com precisão a compatibilidade entre pessoas usando técnicas sofisticadas de aprendizado de máquina e o *crowdsourcing*, que é um método popular de obter rapidamente grandes conjuntos de dados de treinamento para aprendizado de máquina, incluindo análise de sentimentos. Eles encontram confusões de classificação em todos os esquemas, sendo que os esquemas Circumplex e EARL demonstraram maior confusão em uma ampla gama de classes. O esquema das seis emoções básicas foi o mais útil para análise de linguagem emotiva em termos de facilidade de uso e treinamento de algoritmos de aprendizado de máquina

supervisionados. As seis emoções básicas são baseadas na expressão funcional de processos biológicos e evolutivos humanos, e se revelaram em descobertas empíricas de estudos transculturais através do reconhecimento de expressões faciais. São elas: raiva, desgosto, medo, felicidade, tristeza e surpresa (Ekman e Friesen, 1987).

Neste sentido, Johnson et al. (2020) sugere que os designers de plataformas criem espaços para comunidades vulneráveis, permitindo a participação das partes interessadas, mantendo a privacidade e a segurança dos membros mais vulneráveis. Eles partem do diagnóstico de comunidades on-line que tratam de demência. Analisando postagens originais de um período de sete anos em um fórum, destacam que o aprendizado de máquina pode ser útil para maior inclusão das pessoas com deficiência ou demência, até mesmo identificando as mudanças de padrões linguísticos.

Joyce Lau e Agius (2021) propõem um framework MCI-GaTE (MCI-Game Therapy Experience) que pode ser usado para desenvolver jogos como ferramentas eficazes de reabilitação cognitiva e física, e a estrutura de jogo baseada em aprendizado de máquina fornece treinamento para reconhecimento, aprendizado, memória, recordação e outras funções executivas.

Dellerman et al. (2019) se unem as inteligências humana e a artificial ao explicitar que os seres humanos e as máquinas possuem inteligências diferentes que podem ser complementadas, ajudando um ao outro. É sabido que os seres humanos já auxiliam as máquinas, fornecendo assistência em várias partes do processo de aprendizado de máquina para apoiar em tarefas que a inteligência artificial não pode resolver sozinha, fazendo com que, com o uso de aprendizado de máquina, aprenda-se movimentos humanos, ao mesmo tempo que a inteligência humana ensine conhecimentos novos.

Neste ambiente de interação e hibridez, os jogos ganham destaque. A discussão acerca dos afetos, emoções e o campo do aprendizado de máquina tem impactado a área do design de jogos. Dobre e Xueni Pan et al. (2022) analisaram que é um desafio programar Personagens Não-Jogadores (NPCs) para reagir adequadamente a sinais sociais, especialmente em jogos narrativos imersivos em Realidade Virtual (RV). Para isso, utilizou-se de um algoritmo de aprendizado por reforço com recompensas de aprendizado por imitação usando dados brutos e dados derivados socialmente significativos, demonstrando que se pode usar aprendizado de máquina e realidade virtual para projetar experiências mais envolventes.

A partir desta revisão da literatura percebe-se algumas lacunas e limitações. O uso de equipamentos eletrônicos e equipamentos de leitura da atividade cerebral, ainda é limitado em ambientes clínicos devido à precisão e confiabilidade da interface sensorial e algoritmos de tradução dessas tecnologias, que envolvem restrições de tempo, o número de eletrodos e o número de emoções reconhecidas. Há ainda a necessidade de um melhor entendimento de estados afetivos singulares que são quase sempre generalizados pelos sistemas de análise. Ademais, as emoções eliciadas precisam ser analisadas a partir das diferenças entre as que surgem em relação ao evento experimental e o mundo real (Papanastasiou et al., 2020). Essa dificuldade de um olhar mais específico e menos generalista parece estar relacionado ao fato de resultados diagnósticos, por exemplo, serem mais encontrados em grandes bases de dados, onde exames já realizados encontram-se depositados.

Além disso, as amostras nos estudos diagnósticos ainda são limitadas, como uso de grupos muito específicos selecionados para interação com sistemas computadorizados (Forgas-Coll et al., 2021; Antonucci et al., 2021). Quando os estudos são com crianças deve-se compreender que as emoções ainda estão em processo de integração com os outros processos psicológicos como memória e pensamento, o que implica em uma melhor observação dos aspectos emocionais em crianças mais novas. Quando são grupos vulneráveis os estudos encontram dificuldades para a coleta de dados que consigam expressamente revelar diagnósticos.

Nos trabalhos que apresentam uso de *chatbots*, há uma necessidade de construir estrutura de modelagem mais unificada, mais proativa e mais ética, além da possibilidade de conversas mais focadas em objetivos (Zhou et al., 2020). O que se ressalta é que assim como há falta de diretrizes de

design específicas para o desenvolvimento de interfaces de usuário-máquina, o que reforça também a falta de confiança na IA. Isto reforça o argumento de Leimester et al. (2019) que a cocriação de serviços de Inteligência Híbrida entre humanos e agentes inteligentes pode criar um sentimento de aceitação e a confiança.

Portanto, neste trabalho, buscando identificar os objetivos dos estudos que utilizam dos afetos como principal alvo de discussão no campo do aprendizado de máquina, analisar as principais dificuldades e lacunas no âmbito dos estudos sobre afetos em aprendizado de máquina, e identificar os afetos mais estudados, encontra-se que os trabalhos, em geral, lidam diretamente com a afetividade humana ou esta aparece como elemento ligado aos conceitos gerais, como é no caso dos estudos acerca do engajamento. Os trabalhos discutidos se destacam pelo uso de aprendizado de máquina na leitura e análise de dados retirados de resultados de exames diagnósticos ou de redes léxicas, associados a equipamentos de análise da atividade cerebral, destacando o eletroencefalograma (EEG), ou mesmo, o uso de robôs, *chatbots* ou jogos eletrônicos.

Para a questão de pesquisa Q1 ser respondida nesta revisão, sobre como o campo dos afetos tem sido trabalhado no âmbito dos estudos em aprendizado de máquina, nota-se uma evidência em estudos que se utilizam de jogos ou recursos lúdicos para analisar comportamento emocional, eliciar emoções que supram necessidades de comunicação e afeto, treinar e reabilitar pessoas com transtornos do neurodesenvolvimento, associar adversidades na infância e envelhecimento saudável, resistência a preconceitos e discriminações, além de identificar emoções associadas a recente pandemia. O campo dos afetos tem sido explorado no âmbito dos estudos que analisam sentimentos através de algoritmos treinados para identificar sentimentos expressos, como positivo, negativo ou neutro, principalmente para analisar opiniões de consumidores. Já no reconhecimento de emoções, os algoritmos são treinados para identificar as emoções expressas em uma determinada entrada, como uma imagem, um vídeo ou uma fala, aplicados na análise de comportamento humano em situações de avaliação de usuários de serviços de saúde mental. Além disso, a interação humano-máquina também tem sido explorada nesse contexto, com o objetivo de criar sistemas que sejam capazes de reconhecer e responder às emoções dos usuários de forma mais eficiente e personalizada.

Para a questão Q2, compreende-se que o campo dos afetos na aprendizagem de máquina ainda é relativamente novo, e há muitos desafios a serem enfrentados para garantir a precisão e a ética nessas aplicações. Como o treinamento de modelos de aprendizado de máquina depende de grandes quantidades de dados rotulados, a dificuldade começa em conseguir estes dados sensíveis, pois trata-se de avaliações subjetivas das emoções humanas, o que dificulta a quantificação e definição de maneira consistente, especialmente quando tratamos de grupos mais vulneráveis socialmente. Além disso, as emoções são influenciadas por muitos fatores, tais como contexto social e experiências passadas, o que dificulta ainda mais a construção de padrões e métricas consistentes para comparar e avaliar o desempenho de modelos de aprendizado de máquina em relação às emoções. Por fim, a questão ética também esbarra na privacidade e viés algorítmico.

Por fim, para a questão Q3, revela-se que os principais afetos escolhidos para análise são os sentimentos e emoções, personalidade, intenção e comportamento. Emoções expressas ou implícitas, como alegria, tristeza, raiva, medo, surpresa, entre outras, são alvos de análises, assim como analisar intenção de comportamento.

Quanto às limitações deste estudo, percebe-se um foco numa generalização do tema dos afetos no campo do aprendizado de máquina. Como o objetivo do artigo é uma análise geral do campo dos afetos, não há a necessidade de um aprofundamento em um tema específico acerca desta área de aplicação do aprendizado de máquina. Sugere-se, portanto, novas pesquisas que ajudem a esclarecer sobre como o campo dos afetos tem sido trabalhado em âmbitos específicos.

ORCID

Saulo Almeida  <https://orcid.org/0000-0001-7390-3030>

Paulo Caetano da Silva  <https://orcid.org/0000-0002-5038-2460>

CONTRIBUIÇÃO DOS AUTORES

Saulo Almeida: Fundamentação teórica, Metodologia, Análise dos resultados e Redação do rascunho original.

Paulo Da Silva: Análise dos resultados e Redação – revisão.

REFERÊNCIAS

- Andreasson, R., Alenljung, B., Billing, E., & Lowe, R. (2018). Affective Touch in Human – Robot Interaction: Conveying Emotion to the Nao Robot. *International Journal of Social Robotics*, 10(4), 473–491. <https://doi.org/10.1007/s12369-017-0446-3>
- Antonucci, L. A., Raio, A., Pergola, G., Gelao, B., Papalino, M., Rampino, A., Andriola, I., Blasi, G., & Bertolino, A. (2021). Machine learning - based ability to classify psychosis and early stages of disease through parenting and attachment - related variables is associated with social cognition. *BMC Psychology*, 1–15. <https://doi.org/10.1186/s40359-021-00552-3>
- Banziger, T., Pirker, H., & Scherer, K. R. (University of G. (2006). GEMEP - GENEVA Multimodal Emotion Portrayals: A corpus for the study of multimodal emotional expressions. *Proceedings of the Fifth Conference on Language Resources and Evaluation*, 6(May), 15–19.
- Cerri, R. (2017). Aprendizado de máquina: breve introdução e aplicações. *Cadernos de Ciência & Tecnologia, Brasília*, 34(3), 297–313.
- Chae, S. W., & Lee, S. H. (2022). Computers in Human Behavior Sharing emotion while spectating video game play: Exploring Twitch users' emotional change after the outbreak of the COVID-19 pandemic. *Computers in Human Behavior*, 131(January), 107211. <https://doi.org/10.1016/j.chb.2022.107211>
- Cristina, G., Marco, D., & Xueni, G. (2022). Immersive machine learning for social attitude detection in virtual reality narrative games. *Virtual Reality*, 0123456789. <https://doi.org/10.1007/s10055-022-00644-4>
- Dermeval, D., Coelho, J., & Bittencourt, I. (2020). Mapeamento Sistemático e Revisão Sistemática da Literatura em Informática na Educação. In P. A. Jaques, S. Siqueira, I. Bittencourt, & M. Pimentel (Org.), *Metodologia de Pesquisa Científica Em Informática Na Educação: Abordagem Quantitativa* (Livro 2). Sociedade Brasileira de Computação.
- Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid Intelligence. *Business & Information Systems Engineering*, 61, 637–643. <https://doi.org/10.1007/s12599-019-00595-2>
- Domingos, P. (2017). *O Algoritmo Mestre: Como a Busca Pelo Algoritmo de Machine Learning Definitivo Recriará Nosso Mundo* (1th ed.). Novatec Editora.
- Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., LeCompte, W. A., Pitcairn, T., & Ricci-Bitti, P. E. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of personality and social psychology*, 53(4), 712–717. <https://doi.org/10.1037//0022-3514.53.4.712>
- Forgas-coll, S., Huertas-garcia, R., Andriella, A., Alenyà, G., & Huertas-garcia, R. (2023). How do Consumers' Gender and Rational Thinking Affect the Acceptance of Entertainment Social

- Robots? *International Journal of Social Robotics*, 14, 973-994. <https://doi.org/10.1007/s12369-021-00845-y>
- Frewen, P., Schroeter, M. L., Riva, G., Cipresso, P., Fairfield, B., Padulo, C., Haddon, A., Palaniyappan, L., Owolabi, M., Kusi-mensah, K., Polyakova, M., Fehertoi, N., Andrea, W. D., Lowe, L., & Northoff, G. (2020). Neuroscience and Biobehavioral Reviews Neuroimaging the consciousness of self: Review , and conceptual- methodological framework. *Neuroscience and Biobehavioral Reviews*, 112(January), 164–212. <https://doi.org/10.1016/j.neubiorev.2020.01.023>
- Frewen, P., Schroeter, M. L., Riva, G., Cipresso, P., Fairfield, B., Padulo, C., Kemp, A. H., Palaniyappan, L., Owolabi, M., Kusi-Mensah, K., Polyakova, M., Fehertoi, N., D’Andrea, W., Lowe, L., & Northoff, G. (2020). Neuroimaging the consciousness of self: Review, and conceptual-methodological framework. *Neuroscience & Biobehavioral Reviews*, 112, 164–212. <https://doi.org/https://doi.org/10.1016/j.neubiorev.2020.01.023>
- Hu B. (2021). Childhood adversity and healthy ageing: a study of the Chinese older population. *European journal of ageing*, 18(4), 523–535. <https://doi.org/10.1007/s10433-021-00608-8>
- Hutson, J., Taft, J. G., Barocas, S., & Levy, K. (2018). Debiasing Desire: Addressing Bias & Discrimination on Intimate Platforms. *Proceedings of the ACM on Human-Computer Interaction*, 2, 1–18. <https://doi.org/10.1145/3274342>
- Johnson, J., Black, R. W., & Hayes, G. R. (2020). Roles in the Discussion: An Analysis of Social Support in an Online Forum for People with Dementia. *Proceedings of the ACM on Human-Computer Interaction*, 4, 1-30. <https://doi.org/10.1145/3415198>
- Kocoń, J., & Maziarz, M. (2021). Mapping WordNet onto human brain connectome in emotion processing and semantic similarity recognition. *Information Processing and Management*, 58(3), 102530. <https://doi.org/10.1016/j.ipm.2021.102530>
- Lau, S. J., & Agius, H. (2021). A framework and immersive serious game for mild cognitive impairment. *Multimedia Tools and Applications*, 80(20), 31183 - 31237. <https://doi.org/10.1007/s11042-021-11042-4>.
- Marc, W., Rasul, T., Kumar, S., & Ala, M. (2022). Past , present , and future of customer engagement. *Journal of Business Research*, 140(May 2021), 439–458. <https://doi.org/10.1016/j.jbusres.2021.11.014>
- Ocklenburg, S., Packheiser, J., & Hidalgo, G. (2022). Social touch in the age of computational ethology: Embracing as a multidimensional and complex behaviour. *Current Psychology*, 0123456789. <https://doi.org/10.1007/s12144-022-03051-9>
- Papanastasiou, G., Drigas, A., Skianis, C., & Lytras, M. (2020). Heliyon Brain computer interface based applications for training and rehabilitation of students with neurodevelopmental disorders. A literature review. *Heliyon*, 6(May 2019), e04250. <https://doi.org/10.1016/j.heliyon.2020.e04250>
- Petrantonakis, P. C., & Hadjileontiadis, L. J. (2010). Emotion recognition from EEG using higher order crossings. *IEEE Transactions on Information Technology in Biomedicine : A Publication of the IEEE Engineering in Medicine and Biology Society*, 14(2), 186–197. <https://doi.org/10.1109/TITB.2009.2034649>
- Picard, R. W. (1997). *Affective computing*. The MIT Press.
- Plutchik, R. (1997). The circumplex as a general model of the structure of emotions and personality. In R. Plutchik & H. R. Conte (Eds.), *Circumplex models of personality and emotions* (pp. 17–45). American Psychological Association. <https://doi.org/10.1037/10261-001>
- Sætra, H. S. (2020). "The parasitic nature of social AI: Sharing minds with the mindless": Correction. *Integrative Psychological & Behavioral Science*, 54(2), 327. <https://doi.org/10.1007/s12124-020-09536-1>
- Sutoyo, R., Chowanda, A., Kurniati, A., & Wongso, R. (2019). ScienceDirect Designing Designing an an Emotionally Emotionally Realistic Realistic Chatbot Chatbot Framework Framework to to

Enhance Enhance Its Believability with AIML Information States Its Believability with AIML and Information States. *Procedia Computer Science*, 157, 621–628. <https://doi.org/10.1016/j.procs.2019.08.226>

Thanaraj, P., Alex, K., Joseph, N., & Rajangam, V. (2021). Emotion classification from speech signal based on empirical mode decomposition and non-linear features Speech emotion recognition. *Complex & Intelligent Systems*, 7(4), 1919–1934. <https://doi.org/10.1007/s40747-021-00295-z>

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>

Zhou, L., Gao, J., Li, D., & Shum, H.-Y. (2020). The Design and Implementation of XiaoIce, an Empathetic Social Chatbot. *Computational Linguistics*, 46(1), 53–93. https://doi.org/10.1162/coli_a_00368