

| INTELIGÊNCIA ARTIFICIAL

TEMA 1 – PROCESSAMENTO DE LINGUAGEM NATURAL

Além do comportamento inteligente presente em um agente em qualquer situação frente a um problema a ser resolvido, os processos de comunicação com o agente constituem-se num aspecto primordial a ser abordado na área da Inteligência Artificial. A comunicação com os computadores sempre foi feita a partir de linguagens específicas de programação, que definem o que um computador irá fazer, quais recursos irá acessar da memória e de que forma os resultados serão apresentados aos programadores.

O desafio de fazer um agente se comunicar em linguagem natural requereu o estudo, ao longo dos anos, sobre como os seres humanos se comunicam, tema já abordado por filósofos da linguagem e psicólogos. Ainda que a apresentação de resultados de um processamento possa ser feita a partir da criação de frases padrões em um computador, o reconhecimento e a compreensão da fala humana são talvez os maiores dos desafios na concepção de máquinas que tenham a capacidade de comunicação próxima à do ser humano.

A comunicação é a troca intencional de informações provocada pela produção e percepção de sinais extraídos de um sistema compartilhado de sinais convencionais. (...) Em um mundo parcialmente observável, a comunicação pode ajudar os agentes a terem sucesso, porque eles podem aprender informações que serão observadas ou deduzidas por outros. (Russel; Norvig, 2004, p. 765)

O atributo *per se*, que distingue os seres humanos dos animais, é a forma de comunicação por meio de um sistema de troca de mensagens complexas que seguem uma certa estrutura, ou seja, a linguagem. A capacidade de comunicação que o ser humano consegue empreender de forma confiável é alcançada com o trato da linguagem a partir de um número ilimitado de mensagens. O próprio teste proposto por Turing se baseia na definição de uma máquina inteligente a partir da capacidade de comunicação de forma indistinguível daquela entre seres humanos.

Uma das ações a serem executadas por um agente é, portanto, a comunicação. A produção da linguagem consiste em **atos de fala** (Russel; Norvig, 2004). A escrita de qualquer mensagem utilizando sinais específicos de linguagem, em qualquer meio, são atos de fala. Ainda que todos causem mudanças na realidade, a classe de atos de fala denominada *declarativa* tende a produzir um efeito significativo sobre o mundo. No caso de uma declaração sobre um conjunto de coisas presentes no mundo (a realidade relacionada à

mensagem), as declarações tendem a ter um valor de verdade: ou a mensagem está de acordo com a situação da realidade ou a expressão contida na declaração não condiz com o que foi declarado. Em lógica, uma expressão ligada a um estado de coisas no mundo também é denominada de *proposição*. Por exemplo, a declaração *o gato está embaixo da mesa* terá um valor de verdade em correspondência com o mundo se, efetivamente, quem profere a mensagem o faz para expressar uma situação da realidade, em que o gato está efetivamente embaixo da mesa (Figura 1).

Figura 1 – Situação de realidade para a mensagem *o gato está embaixo da mesa*



Fonte: Benny Marty/Shutterstock.

Dessa forma, pode-se constatar que existe um **valor de verdade** na relação da mensagem com o mundo. Caso fosse proferida a expressão *o gato está em cima da mesa*, o valor de verdade a ser atribuído seria falso. Assim, o mundo da linguagem pode ser permeado de expressões que podem ser verdadeiras ou falsas, a respeito dessa com a correspondência do mundo.

A própria linguagem estaria no âmago do processo de representação de conhecimento que o ser humano em sua mente faz a respeito da realidade. Certas ideias são melhor representadas em imagens e outras em palavras, ainda que nem uma nem outra possam reter todas as características do objeto que está sendo representado. As palavras tendem a ser representações simbólicas daquilo que está sendo representado, seja um objeto que é nomeado ou uma proposição a respeito das coisas do mundo (Sternberg, 2014).

O estudo da lógica relacionada com a linguagem teve início ainda no século XIX com o fundador da lógica matemática moderna, Gottlob Frege. Este visava a separação do pensamento lógico, relacionado com as coisas do mundo, de um componente mais psicológico. Ou seja, a distinção do que pode ser considerado como um pensamento com componentes **objetivos**, em contrapartida aos aspectos **subjetivos** presentes no pensamento (Frege, 2013).

A partir do uso da linguagem por parte de um agente, a tarefa-chave é decidir quando um ato de fala será necessário e que ato de fala deve ser o correto. Tem-se então o problema da **compreensão** dos atos de fala (Russel; Norvig, 2004): a partir de um conjunto de entradas ambíguas, o agente deve constatar qual é a situação do mundo e então decidir por um curso de ação.

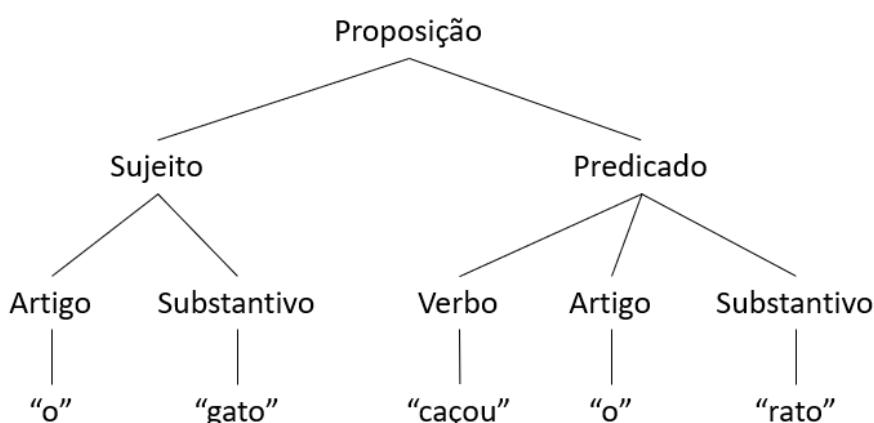
Na comunicação entre dois agentes, um emissor (E) e outro receptor (R), pode ser decomposta em sete processos (Russel; Norvig, 2004):

1. Intenção: o emissor E decide que há uma proposição P que vale a pena informar para o receptor R.
2. Geração: o emissor E planeja transformar a proposição P em uma expressão vocal V que possa ser processada por R e compreendida com relação ao seu significado.
3. Síntese: o emissor E produz uma realização física V, ou seja, utiliza um meio para o envio da expressão V para R.
4. Percepção: R percebe a realização física V decodificando em uma expressão W (reconhecimento de fala).
5. Análise: R obtém significados possíveis de W, perfazendo a análise sintática, a interpretação semântica e a interpretação pragmática.
 - 5.1. Análise sintática: é o processo de construção de uma árvore de análise, onde os nós desta árvore representam as palavras.
 - 5.2. Interpretação semântica: é o processo de extração do significado de uma expressão vocal em alguma linguagem de representação.
 - 5.3. Interpretação pragmática: é o processo que relaciona a expressão vocal com o contexto em que é articulada.
6. Eliminação da ambiguidade: o receptor R deduz que o emissor E pretendia transmitir a proposição P.
7. Incorporação: o receptor R decide acreditar na proposição P, adicionando ao seu conjunto de representações como um conhecimento adicional sobre o mundo.

Portanto, o **processamento de linguagem natural** se refere ao desenvolvimento de modelos computacionais para a execução de tarefas que dependem de informações que estejam expressas em alguma língua natural. A compreensão da linguagem exige a compreensão tanto do assunto quanto do contexto, não somente o entendimento da estrutura das frases (Russel; Norvig, 2004).

Para a compreensão de uma frase, um agente precisa alcançar o significado presente nessa frase. Dessa forma, a partir da análise de uma frase ou proposição, é possível ao agente obter um conhecimento do que está sendo expresso, ou seja, o agente visa alcançar a interpretação **semântica** contida em uma proposição.

Figura 2 – Exemplo de uma árvore sintática para a frase *o gato caçou o rato*



Uma **gramática** é um conjunto finito de regras que especificam uma linguagem. As linguagens formais, como as linguagens de computação, sempre contêm uma gramática oficial, que é especificada em algum manual ou livro. Já as linguagens naturais não possuem uma gramática oficial. No entanto, os linguistas buscam descobrir as propriedades utilizando uma metodologia científica, incorporando em alguma espécie de gramática (Russel; Norvig, 2004). A forma mais simples de determinar uma gramática é especificar um conjunto de símbolos terminais (palavras da linguagem), símbolos não terminais (componentes das sentenças) e um conjunto de regras de produção (regras se/então) (Pereira, s/a).

Por exemplo, a seguinte gramática permite definir um pequeno fragmento da língua portuguesa:

```
proposição -> sujeito predicado
sujeito -> artigo substantivo
predicado -> verbo artigo substantivo
artigo -> o
substantivo -> gato | rato
verbo -> caçou
```

Analizando-se esta gramática, os símbolos terminais são *o*, *gato*, *rato* e *caçou*. O conjunto das regras que exibem os símbolos terminais são denominados de *léxico*, ou seja, o conjunto de palavras permitidas. A regra de produção “proposição → sujeito predicado” indica que uma proposição é composta de um sujeito e de um predicado. A regra *substantivo* → *gato* | *rato* define que um substantivo pode assumir o conteúdo como sendo *gato* ou *rato*. Como símbolos não terminais, há os antecedentes das três primeiras regras (proposição, sujeito e predicado).

A **análise sintática** é, portanto, o processo de se encontrar uma árvore de análise para uma dada cadeia de entrada e uma gramática em questão, como na frase *o gato caçou o rato produz* que, por meio da gramática, foi definida anteriormente (Figura 2). Ela pode ser vista como um processo de busca de uma árvore de análise (Russel; Norvig, 2004). Um analisador sintático, denominado de *forma comum de parser*, executa o processo de construir uma árvore de representação da frase analisada para que seja possível a interpretação semântica no passo seguinte.

TEMA 2 – EXTRAÇÃO DE DADOS DE TEXTO

A mineração de dados de texto é um processo que visa a descoberta de conhecimento que utiliza técnicas de extração de dados a partir de textos. Para isso, leva em consideração o uso de algoritmos que processam textos e identificam informações úteis que não poderiam emergir utilizando-se métodos tradicionais de consulta. Os dados de texto seguem uma forma de armazenamento diferenciada, sendo eminentemente mantidos em bases de dados não estruturadas (Morais; Ambrósio, 2007).

Conforme Barion e Lago (2008, p. 125), “a mineração de textos é um conjunto de métodos utilizados para navegar, organizar, achar e descobrir informações em bases de textos”. Pode-se considerar como um elemento dos processos de *data mining*, visando especificamente textos. Conforme Passos e

Aranha (2006), a área da mineração de textos se caracteriza por ser interdisciplinar, envolvendo outras áreas de conhecimento como informática, estatística, linguística e ciências cognitivas.

(...) A mineração de textos define técnicas de extração de padrões ou tendências de grandes volumes de textos em linguagem natural, normalmente, para objetivos específicos. Inspirado pelo data mining ou mineração de dados, que procura descobrir padrões emergentes de bancos de dados estruturados, a mineração de textos pretende extrair conhecimentos úteis de dados não estruturados ou semiestruturados. (Barion; Lago, 2008, p. 125)

As etapas do processo de mineração de textos (Figura 3) são: seleção de documentos (também denominada de *criação de corpus*), definição do tipo e abordagem dos dados (análise estatística ou semântica), preparação dos dados, indexação e normalização, cálculo da relevância de termos, seleção dos termos e pós-processamento ou análise dos resultados (Morais; Ambrósio, 2007).

A **análise semântica** emprega técnicas baseadas em processamento de linguagem natural, que buscam avaliar a sequência de termos dentro do contexto dos textos, buscando identificar a sua função. Técnicas de análise semântica tentam identificar a importância das palavras dentro da estrutura de suas orações. Entretanto, quando se utiliza um texto único, algumas funções podem ser identificadas adotando-se um grau de importância. É fundamentada basicamente em processamento de linguagem natural (Morais; Ambrósio, 2007).

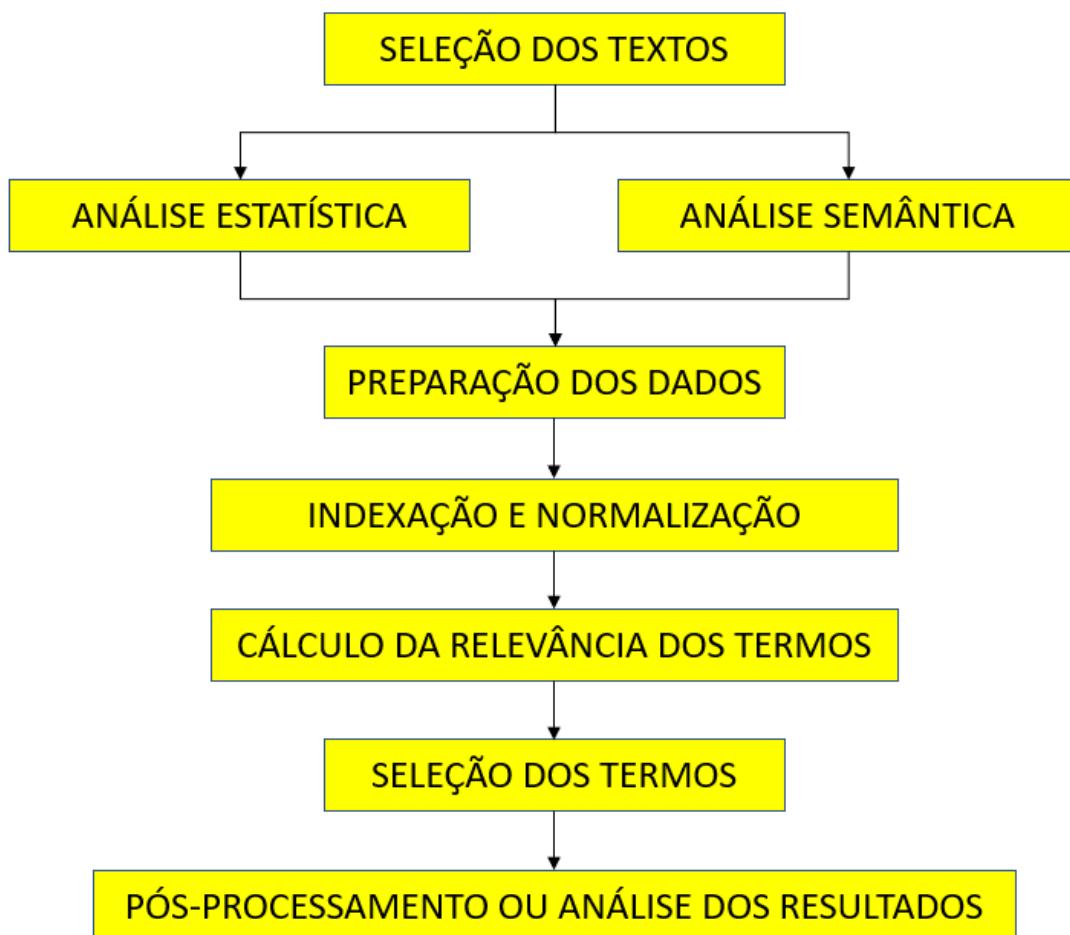
A **análise estatística** considera a quantidade de vezes que um termo aparece no dentro do texto. A noção geral é a existência de um aprendizado estatístico a partir dos dados que incluem as fases de codificação, estimativa e modelos de representação de documentos (Morais; Ambrósio, 2007).

A **preparação dos dados** envolve a seleção dos dados de interesse e o trabalho inicial para tentar identificar o núcleo que irá melhor representar o conteúdo dos textos. Além de permitir uma redução dimensional, o objetivo é tentar encontrar similaridades conforme a morfologia ou o significado dos termos dentro dos documentos.

Por meio da **indexação**, as palavras presentes no texto são armazenadas em uma estrutura de índices para que se viabilize a pesquisa de documentos por meio das palavras que ele contém. Arquivos invertidos são geralmente utilizados para a implementação de índices ordenados, que pode ser visualizado como uma lista ordenada de palavras contendo um ponteiro para cada um dos textos nos quais ela ocorre. Tais tipos de índices invertidos são usados para melhoria do desempenho e a funcionalidade das buscas (Barion; Lago, 2008).

O processo de indexação é composto de análise léxica, remoção de *stopwords*, *stemming*, seleção dos termos-índice, determinação de pesos e criação de tesouros.

Figura 3 – Etapas do processo de mineração de textos



Fonte: Adaptado de Morais e Ambrósio, 2007.

Quadro 1 – Etapas de indexação

ITEM DE INDEXAÇÃO	DESCRIÇÃO
Análise Léxica	Etapa para converter uma sequência de caracteres em uma sequência de palavras candidatas a termos do índice.
Remoção de <i>stopwords</i>	Exclusão de palavras que sejam preposições, artigos, conjunções, verbos, nomes, advérbios e adjetivos.
<i>Stemming</i>	Remoção das variações das palavras para extração do radical.

Seleção dos termos-índice	Determinação de quais palavras ou radicais serão utilizados como elementos de indexação.
Determinação de pesos	Uso de medidas de frequência relativa (peso do termo), calculados proporcionalmente à frequência de ocorrência de cada termo em cada documento.
Criação de tesouros	Refinamentos que consistem em associação de termos, representando um vocabulário controlado, pertencendo a um domínio específico de conhecimento ou mais genérico para uma língua.

Fonte: Adaptado de Barion e Lago, 2008.

Em algumas situações, é necessário fazer a comparação de textos, visando, por exemplo, a categorização dos textos produzidos. A comparação pode contar com algumas métricas para indicar a distância que um texto apresenta em relação a outro. Isto pode ser importante inclusive para análises de aglomeração, em algoritmos como *k-means*, para a classificação *a priori* de conjuntos de textos.

Uma métrica comumente utilizada é a distância de Levenshtein, a qual se utiliza de operações necessárias para tornar um texto igual a outro. Por exemplo, a distância entre *casa* e *rasa* é um; a distância entre *poste* e *teste* é dois (Amaral, 2016).

Outra métrica utilizada é a distância de Jaro. Nesta, cada letra ou caractere de um termo ou texto é comparado com cada letra ou caractere do segundo termo ou texto. O objetivo é verificar quais caracteres são iguais e também levando em consideração a ordem em que eles se encontram. Por exemplo, *canto* e *conta* possuem os mesmos caracteres, porém as vogais estão em ordem diferente. Uma variação desta métrica é a distância de Jaro-Winkler, que favorece a comparação do começo dos termos até uma parte definida no interior deles.

TEMA 3 – ANÁLISE DE SENTIMENTOS

Conforme Luo et al. (2013), a **análise de sentimentos** (também chamada de *mineração de opinião*) refere-se à aplicação do processamento de linguagem natural, linguística computacional e análise de texto para identificar e classificar opiniões subjetivas nos materiais de origem (um documento ou uma frase, por exemplo). De um modo geral, a análise de sentimentos visa determinar a atitude de um escritor em relação a algum tópico ou a polaridade contextual geral de um

documento. A atitude pode ser seu julgamento ou avaliação, estado afetivo (isto é, o estado emocional do autor ao escrever) ou a comunicação emocional pretendida (isto é, o efeito emocional que o autor deseja ter no leitor).

A análise de sentimentos reúne várias áreas de pesquisa, como processamento de linguagem natural, mineração de dados e mineração de texto, e está rapidamente se tornando de grande importância para as organizações, que se esforçam para integrar métodos de inteligência computacional em suas operações e tentam esclarecer e melhorar seus produtos e serviços (Farhadloo; Rolland, 2016). Assim, o objetivo é descobrir as opiniões das pessoas expressas na linguagem escrita (texto). O termo *sentimento* significa “o que alguém sente por algo”, “experiência pessoal, sentimento próprio”, “uma atitude em relação a algo” ou “uma opinião”.

Geralmente, a análise de sentimentos classifica expressões de texto nos materiais de origem em dois tipos:

- 1) fatos (objetivo): expressões objetivas sobre entidades, eventos e seus atributos, por exemplo, *comprei um iPhone ontem*;
- 2) opiniões (subjetivas): expressões subjetivas de sentimentos, atitudes, emoções, avaliações ou sentimentos em relação a entidades, eventos e seus atributos, por exemplo, *eu realmente amo essa nova câmera*.

Cabe ressaltar que nem todas as frases subjetivas contêm opiniões, por exemplo, *quero um telefone com boa qualidade de voz*; e nem todas as frases objetivas não contêm opiniões, por exemplo, *o fone de ouvido quebrou em apenas dois dias!* (Luo et al., 2013).

No exemplo acima, é claro que a sentença (1) seria identificada como um fato; e as frases (2), (3) e (4) como opiniões. No entanto, para as frases (5), (6) e (7), pode ser difícil determinar automaticamente as expressões de opinião no iPhone. Geralmente, as opiniões têm alvos (objetos e seus atributos) nos quais opiniões são expressas. A seguir, é mostrada uma categorização para descrever formalmente as opiniões (Luo et al., 2013).

- Objeto: uma entidade que pode ser um produto, serviço, indivíduo, organização, evento ou tópico, por exemplo, iPhone.
- Atributo: um objeto geralmente possui dois tipos de atributos: (1) componentes, por exemplo, bateria, teclado/tela sensível ao toque; e (2) propriedades, por exemplo, tamanho, peso, cor, qualidade da voz.

- Atributos explícitos e implícitos: atributos explícitos refere-se àqueles que aparecem na opinião, por exemplo, *a duração da bateria deste telefone é muito curta*; e atributos implícitos se referem àqueles que não aparecem na opinião, por exemplo, *este telefone é muito grande* (no tamanho do atributo).
- Detentor de opinião: a pessoa ou organização que expressa a opinião.
- Orientação de opinião: isto é, polaridade, por exemplo, positiva, negativa ou neutra.
- Força da opinião: nível/escala/intensidade da opinião indicando quanto forte é, por exemplo, contente, feliz, alegre e extática, cuja força é incremental.
- Opinião: uma pessoa ou organização que expressa um sentimento positivo ou negativo sobre um atributo específico de um objeto em um determinado momento. Isso pode ser representado como uma quíntupla: <objeto, atributo, orientação, formador de opinião, tempo>.

Com base nas técnicas de aprendizado de máquina, abordagens de aprendizado supervisionadas são usadas para aplicar na classificação de orientação de sentimentos. A ideia básica é alavancar técnicas de aprendizado supervisionado para encontrar padrões em exemplos conhecidos e aplicá-los a novos documentos, para classificar automaticamente a orientação de sentimentos de novos documentos. Assim, o objetivo do aprendizado supervisionado é treinar e obter um classificador de opinião que contenha algumas classes de opinião-alvo, por exemplo, positivas *versus* negativas. A maioria das abordagens de aprendizado supervisionado existentes pode ser usada para atingir a meta (ou seja, obter o classificador de opinião). Agora, os métodos populares de aprendizado supervisionado incluem principalmente (Luo et al., 2013):

- Naïve Bayes (NB): um classificador probabilístico simples baseado na aplicação do teorema de Bayes com fortes suposições de independência (ingênuas).
- Entropia Máxima (EM): um modelo probabilístico que estima a distribuição condicional do rótulo da classe.
- Máquinas de vetores de suporte (SVM): uma representação dos exemplos como pontos no espaço em que os vetores de suporte são calculados para fornecer uma melhor divisão de pontos/exemplos em categorias.

- Modelo de Regressão Logística (LR): um modelo de LR prevê as classes a partir de um conjunto de variáveis que podem ser contínuas, discretas ou uma mistura.

TEMA 4 – CHATBOTS E ASSISTENTES VIRTUAIS

As formas de interação entre homem e máquina evoluíram significativamente nos últimos anos, incluindo a educação, que passou por mudanças substanciais em suas práticas devido ao surgimento de novas tecnologias, seguindo caminhos diferentes, que levantam novas possibilidades diante de inúmeras novas demandas. Entre várias inovações, o uso de tutores inteligentes é uma tendência que reforça as crescentes necessidades de maior interatividade. As tecnologias que oferecem interfaces de linguagem natural são uma boa opção para melhorar as interações por meio de agentes de interface de conversação incorporados, ou dos chamados *chatterbots*. Os alunos poderiam ter apoio adicional de tutores de IA comunicando-se em linguagem natural e aprendendo conteúdos com base em conceitos, exemplos, ilustrações ou outros links, gerenciados por um assistente cognitivo (Medeiros et al., 2019).

Um aspecto relevante a ser discutido é a nomeação desse sistema. O termo *assistente cognitivo* é abrangente, porque o sistema não apenas retorna respostas para uma determinada pergunta, mas também aprende com suas interações com os alunos. Muitas publicações de pesquisa nesse campo fornecem um conjunto de denominações diferentes de *assistente cognitivo* ou *tutor cognitivo*, mas também podemos encontrar termos como *assistente de chatbot*, *assistente pessoal*, *tutores inteligentes*, *assistente de exploração de conhecimento*, *assistente de aprendizagem ontológica* e *agente conversacional* (Medeiros et al., 2019).

A construção de sistemas que se comportam como seres humanos é um dos objetivos de pesquisa no campo da Inteligência Artificial. No campo dos sistemas que interagem com o ser humano por meio de linguagem natural, uma das primeiras tentativas de simulação de um especialista humano foi o *software Eliza*, que tentava simular o atendimento de um psicólogo rogeriano a um paciente humano (Weizenbaum, 1976).

Atualmente, o surgimento de sistemas de perguntas e respostas (*Questions and Answers*, ou Q&A) está mudando a análise de dados, principalmente para aprendizado de máquina e tarefas de *Big Data*. As perguntas e respostas

consistem em uma forma especial de recuperação de informações que visa recuperar o conhecimento *on-line*, unificando a linguagem natural e a representação do conhecimento, inferência lógica e pesquisa semântica. Um exemplo é a inovadora tecnologia do IBM Watson, que mostra a relevância de tais sistemas diante do atual crescimento maciço de dados e informações (Medeiros et al., 2019).

O IBM Watson foi desenvolvido baseado na arquitetura Deep Q&A. Tornou-se famoso devido à participação do programa de televisão americano *Jeopardy*. O sistema Watson apresentou sua performance no mesmo nível de um especialista humano em termos de precisão, confiança e velocidade requerida pelo programa de televisão. Pode-se dizer que ele abriu caminho para os denominados *sistemas cognitivos*. Uma série de outros assistentes cognitivos foram desenvolvidos e estão à disposição do mercado:

- SIRI (Apple);
- Cortana (Microsoft);
- Assistant (Android);
- ROSS (Lawyer Assistant);
- Google Now;
- Alexa (Amazon);
- Amy/Andrew;
- Shae (PH360);
- Otto (Samsung).

Na área de serviços para o cliente, há um forte interesse das empresas nos *chatbots*, no sentido de proporcionar um atendimento ao mesmo tempo personalizado e amplo. Áreas como telefonia, fornecimento de energia, serviços e pós-venda podem se beneficiar com relação ao uso de *chatbots* e assim diminuir os custos em geral atribuídos ao grande efetivo de pessoal envolvido nessas atividades.

De maneira geral, em um contexto de prestação de serviços a clientes, pode-se pensar em um *chatbot* contendo três partes (Accenture, 2016):

- Interface (*User Interface*, ou UI): o aplicativo utilizado para o contato com o cliente, seja uma sessão de *chat* ou um aplicativo de mensagens. As interfaces de usuário podem ser dotadas de síntese de fala, tal como Siri, Alexa, Cortana etc.

- Inteligência (*Artificial Intelligence*, ou AI): é a inteligência do *chatbot* propriamente dita, a qual permite a compreensão dos textos provenientes do usuário em linguagem natural. A maioria das soluções de problemas em um *chatbot* é feita por meio de regras predefinidas em uma árvore de decisão, ou mesmo dotado de aprendizagem de máquina para aprender pela interação com o usuário.
- Integração (*Systems Integration*, ou SI): o *chatbot* pode ficar integrado a outros sistemas ou plataformas, como o ERP ou o CRM da empresa, permitindo um alcance mais efetivo de comunicação com o cliente a partir das informações obtidas na interação da empresa com o cliente.

Entretanto, o escopo de atuação dos *chatbots* não são bons para todos os casos. *Chatbots* são efetivos quando as requisições de um cliente acontecem em uma área específica, na qual as soluções são bem conhecidas e predizíveis, tais como resolução de problemas ou casos de educação de clientes nos quais existe um *script* a ser seguido. No caso de um contexto de negócio em que os clientes estejam estressados, como no caso de fazer uma reclamação, agentes humanos são melhores para fazer essa abordagem.

Em situações em que existam múltiplos passos ou parâmetros de entrada, um *chatbot* pode se comunicar utilizando linguagem natural para compreender a solicitação de um cliente, de maneira a executar a tarefa de maneira rápida. Por exemplo, no reagendamento de uma viagem, o cliente pode perguntar: *você pode reservar um vôo para São Paulo para a manhã do dia 10 de março, a partir das 3 horas, com o assento na janela?*

Os *chatbots* também podem ser bem-vindos para os casos de uso frequente em que as respostas são muito similares para diferentes tipos de clientes e precisam ser dadas para um volume massivo de clientes. Dessa forma, pode-se criar um modelo de negócios no qual um *chatbot* automatize a maioria dos atendimentos, deixando os agentes humanos livres para atender aos casos mais complexos (Accenture, 2016).

A Figura 4 mostra um exemplo de interação com o assistente cognitivo THOTH, criado para revisão de conteúdos em disciplinas de cursos superiores. (Medeiros et al., 2019).

Figura 4 – Exemplo de conversação com o assistente cognitivo THOTH

```
(connecting...)
[Thoth]: Hello, User! How are you?
[Thoth]: Welcome!
[Thoth]: In the last session, you asked 'what does artificial neuron mean'...
[Thoth]: Your last object viewed was 'artificial neuron'...
[User]: Thanks!
[Thoth]: You're welcome!
[User]: Is there a concept about neural networks?
[Thoth]: (Searching: neural network, concept)
[Thoth]: According to Haykin (2004), neural network means a processor distributed massively and in parallel, consisting of simple processing units, having the natural inclination to store experimental knowledge and make it available to use.
[User]: What's my overall performance?
[Thoth]: (Searching: help, overall performance)
[Thoth]: Your overall performance on the exercises is 71,93%.
[User]: What contents do I need to reinforce?
[Thoth]: (Searching: help, reinforcement of objects)
[Thoth]: Objects that you need to review are: act rationally (50%), think rationally (52%), rule (55%), lines of research (55%) and fact (57%)
[Thoth]: Would you like to learn about another object?
[User]: Give me a link about neural networks
[Thoth]: (Searching: neural network, link)
[Thoth]: Ok... from Wikipedia, I have the following link: Neural Network (https://en.wikipedia.org/wiki/Artifical\_neural\_network)
[Thoth]: Do you want to visit this site?
[User]: Yes
[Thoth]: I can't open this link!
[Thoth]: I'm letting the tutor know, and he will check it soon!
```

TEMA 5 – SIMULADORES COMPUTACIONAIS

A simulação é um conceito de grande utilidade e essencial para o desenvolvimento científico e tecnológico atualmente. Ao longo do tempo, adquiriu importância como área de estudo a partir das possibilidades proporcionadas pelo desenvolvimento da computação. Conforme Pidd (1998), a simulação computacional consiste no uso de um modelo como base para exploração e experimentação da realidade (Medeiros et al., 2014).

Simon (1981) afirmou que se trata de uma técnica para aquisição de compreensão e predição do comportamento de sistemas, antecedendo o computador digital, podendo assumir a forma de um experimento de pensamento que na verdade nunca foi implementado dinamicamente.

A simulação permite ao analista realizar estudos sobre os correspondentes sistemas modelados para responder questões do tipo “O que aconteceria se?”. O principal apelo ao uso dessa ferramenta é que tais questões podem ser respondidas sem que os sistemas sob investigação sofram qualquer perturbação, uma vez que os estudos são realizados no computador (Freitas Filho, 2008, p. 23).

A partir do exposto, é interessante assinalar que a simulação computacional pode ser realizada mesmo para sistemas que ainda não existem,

permitindo antecipar problemas que, somente após a realização física dos mesmos, teriam a possibilidade de serem encontrados. Portanto, a simulação como método de pesquisa exige uma abordagem quanto à origem do conhecimento, tanto racional quanto empírica. Racional por envolver o esforço intelectual na modelagem conceitual dos problemas (com base no objeto concreto ou produtos de experimentos de pensamento); e empírica por permitir um processo de depuração dos problemas abordados baseado num método de tentativa e erro (Medeiros et al., 2014).

As bases da simulação como um método ou técnica antecedem o próprio estudo da computação e sua fundamentação pode ser encontrada a partir da teoria dos sistemas intencionais de Daniel Dennett. Ao longo das eras, a intencionalidade dos seres vivos foi se tornando mais e mais complexa a partir da perspectiva das criaturas mais simples. Dennett (1997) descreve o que foi por ele denominado de *torre de gerar e testar*, como um processo de simulação do ambiente externo que emergiu a partir da deriva dos processos de evolução. Dessa forma, os seres vivos seriam divididos em (Medeiros et al., 2014):

- Criaturas darwinianas: com alusão à teoria da evolução de Darwin, pelo processo de seleção natural há a sobrevivência dos organismos mais aptos ao seu ambiente externo, a partir de uma grande quantidade de candidatos. As informações genéticas dos sobreviventes são propagadas aos seus descendentes.
- Criaturas skinnerianas: estando de acordo com a teoria comportamentalista de Skinner, o organismo usa a tentativa e erro no fornecimento de respostas de forma cega, até que uma delas é selecionada por reforço, sendo por sua vez memorizada e recuperada com maior probabilidade para as próximas respostas a serem transmitidas a partir dos mesmos ou semelhantes estímulos. As chances de sobrevivência dependem do organismo fornecer a resposta correta em grande parte das situações.
- Criaturas popperianas: conforme Popper, existe um ambiente interno no organismo que simula o ambiente externo, responsável pelo fornecimento das respostas após serem testadas neste mesmo ambiente interno. Com a presença desse meio interno seletivo, o organismo consegue sobreviver de maneira mais vantajosa, pois é capaz de visualizar de forma antecipada as respostas possíveis, atuando de maneira bem melhor do que fornecer respostas por acaso, no caso das criaturas darwinianas, ou por respostas

semelhantes a uma anteriormente dada, no caso das criaturas skinnerianas.

- Criaturas gregoriana: conforme os estudos de Richard Gregory (1981), Dennett argumenta que os artefatos mais bem projetados pelo ser humano não são apenas resultados da inteligência, mas sim dotadores de inteligência, permitindo que se manifeste uma espécie de inteligência potencial externa. No alvorecer da raça humana, o uso de ferramentas é associado a um grande aumento da inteligência. Quanto mais bem projetada for a ferramenta (ou seja, quanto mais informação está embutida nela), maior será o potencial de inteligência que ela irá conferir ao seu usuário. Dentre as ferramentas mais importantes, há as ferramentas mentais, as palavras. As palavras e outras ferramentas mentais têm o poder de conferir a uma criatura gregoriana um meio interno que lhe permite construir geradores e testadores de movimento cada vez mais elaborados.

Dessa forma, o ser humano é, na classificação de Dennett, uma criatura gregoriana que possui um ambiente interno que atua como um simulador da realidade. O desenvolvimento de uma estrutura tal como o cérebro permitiu o surgimento de criaturas que pudessem reter em seu ambiente interno as diferentes situações que poderiam ser enfrentadas no ambiente externo e, assim, garantir uma chance cada vez maior de sobrevivência. Sobreviver, nesse caso, não apenas de forma reativa, mas também proativa, no sentido de modificar o próprio ambiente externo e torná-lo mais propício ao surgimento de estruturas sociais mais estáveis (Medeiros et al., 2014).

Por essa exposição, a simulação computacional pode ser entendida como uma extensão do processo de simulação que o próprio cérebro faz da realidade externa. Posiciona-se como uma ferramenta, racional e empírica, altamente elaborada à disposição do ser humano para expandir ainda mais as possibilidades de conhecimento da própria realidade concreta e também de realidades possíveis ou mesmo sobre situações hipotéticas ou imaginárias, fruto do pensamento humano (Medeiros et al., 2014).

Law e Kelton (1991) apresentam as seguintes **vantagens** para a utilização da modelagem e simulação computacional em estudo de sistemas, permitindo:

- a) a abordagem de sistemas complexos com elementos estocásticos sem condições de tratamento por técnicas analíticas;

-
- b) um controle melhor sobre as condições experimentais em relação à experimentação no sistema real;
 - c) a experimentação de forma interativa;
 - d) a replicação precisa dos experimentos e o teste de diferentes cenários para o sistema;
 - e) a simulação de longos períodos em um tempo menor;
 - f) maior economia do que testes envolvendo sistemas reais.

Como **desvantagens** para a utilização de simulação computacional em estudo de sistemas, Law e Kelton (1991) apontam que:

- a) a simulação depende da validade do modelo desenvolvido. Caso o modelo concebido não represente de forma fidedigna o sistema ou se os dados de entrada não são confiáveis, não haverá respostas adequadas para a solução do problema;
- b) a técnica da simulação não é propriamente uma técnica de otimização, sendo possível somente o teste de alternativas fornecidas pelo usuário;
- c) estudos de simulação podem ser demorados e consumir um alto nível de recursos.

REFERÊNCIAS

- ACCENTURE INTERACTIVE. **Chatbots in Customer Service**. Disponível em: <https://www.accenture.com/t00010101T000000_w__/br-pt/_acnmedia/PDF-45/Accenture-Chatbots-Customer-Service.pdf>. Acesso em: 10 fev. 2020.
- AMARAL, F. **Introdução à ciência de dados**: mineração de dados e Big Data. Rio de Janeiro: Alta Books, 2016.
- BARION, E. C. N.; LAGO, D. Mineração de textos. **Revista de Ciências Exatas e Tecnologia**, v. III, n. 3, p. 123-140, 2008. Disponível em: <<https://seer.pgsskroton.com/index.php/rcext/article/viewFile/2372/2276>>. Acesso em: 1 fev. 2020.
- DENNETT, D. C. **Tipos de mentes**: rumo a uma compreensão da consciência. Rio de Janeiro: Rocco, 1997.
- FARHADLOO, M.; ROLLAND, E. **Fundamentals of Sentiment Analysis and Its Applications**. In: PEDRYCZ, W.; CHEN, S. M. **Sentiment Analysis and Ontology Engineering. Studies in Computational Intelligence**, v. 639. Springer, 2016.
- FREGE, G. **Ensayos de Semántica y Filosofía de la Lógica**. Madrid-España: Tecnos, 2013.
- FREITAS FILHO, P. J. **Introdução à modelagem e simulação de sistemas com aplicações em ARENA**. 2. ed. Florianópolis: Visualbooks, 2008.
- GREGORY, R. **Mind in science**: a history of explanations in psychology and physics. Londres: Weidenfeld & Nicolson, 1981.
- LAW, A. M.; KELTON, W. D. **Simulation Modeling & Analysis**. Nova Iorque: McGraw-Hill Books, 1991.
- LUO, T. et al. Sentiment Analysis. In: **Trust-based Collective View Prediction**. Nova Iorque: Springer, 2013.
- MEDEIROS, L. F.; MOSER, A.; KOLBE JUNIOR, A.; A Cognitive assistant that uses small talk in tutoring conversation. **International Journal: Emerging Technologies in Learning**, v.14, p. 138-159, 2019.
- _____.; SANTOS, N. Simulação computacional como técnica de pesquisa na administração. **Intersaber (Facinter)**, v. 9, p. 463-485, 2014.

MORAIS, P. A. M.; AMBRÓSIO, A. P. L. **Mineração de Textos**. Relatório Técnico. INF 005/07, dez. 2007. Disponível em: <http://www.portal.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_005-07.pdf>. Acesso em: 1 fev. 2020.

PASSOS, E.; ARANHA, C. A tecnologia de mineração de textos. **Revista Eletrônica de Sistemas de Informação**, v. 5, n. 2, 2006. Disponível em: <<http://www.periodicosibepes.org.br/index.php/reinfo/article/view/171>>. Acesso em: 1 fev. 2020.

PEREIRA, S. L. **Processamento de Linguagem Natural**. Disponível em: <<https://www.ime.usp.br/~slago/IA-pln.pdf>>. Acesso em: 15 jan. 2020.

PIDD, M. **Modelagem empresarial**: ferramentas para tomada de decisão. Porto Alegre: Artes Médicas, 1998.

RUSSEL, S.; NORVIG, P. **Inteligência Artificial**. Rio de Janeiro: Campus, 2004.

SIMON, H. **The sciences of the artificial**. 2. ed. Cambridge: The MIT Press, 1981.