



# FERRAMENTAS TECNOLÓGICAS DE MACHINE LEARNING E OS PROCESSOS DO CICLO DO CONHECIMENTO, UM ESTUDO EXPLORATÓRIO

Arthur Anderson Pesco<sup>1</sup>; Iara Carnevale de Almeida<sup>2</sup>

<sup>1</sup>Acadêmico do Curso de Engenharia de Software, UNICESUMAR, Maringá-PR. Bolsista do Programa de Iniciação Científica do Instituto Cesumar de Ciência, Tecnologia e Inovação PIC/ICETI.

<sup>2</sup>Orientadora, Doutora, Bolsista do Programa Produtividade em Pesquisa do Instituto Cesumar de Ciência, Tecnologia e Inovação - ICETI. Programa de Pós-Graduação em Gestão do Conhecimento da UNICESUMAR, Maringá/PR.

**RESUMO:** O presente artigo descreve o ciclo do conhecimento nas ferramentas tecnológicas disponibilizadas pela Inteligência Artificial. A área de Inteligência Artificial tem sido, cada vez mais, reconhecida como uma área em franco desenvolvimento. Diversas são as suas subáreas, tais como aprendizagem de máquina, redes neurais, computação cognitiva e aprendizado profundo. Todas estas trabalham com considerável volume de dados, onde um dos grandes objetivos é o de gerar novo conhecimento. Como resultado deste estudo, pretende-se determinar o ciclo do conhecimento em cada uma das ferramentas tecnológicas, de forma a permitir uma melhor compreensão deste para a Gestão do Conhecimento nas organizações.

**PALAVRAS-CHAVE:** Inteligência Artificial, *Machine Learning*, Gestão do Conhecimento

## 1 INTRODUÇÃO

Inteligência artificial (IA) é um conceito que abrange diferentes metodologias que são capazes de criar máquinas inteligentes. Uma das subáreas da IA que tem se destacado nos últimos anos é a de aprendizado de máquina (AM ou *Machine Learning*). Segundo Mohri et. al (2012), esta subárea explora métodos computacionais que utilizam a experiência para melhorar o desempenho e fazer previsões mais precisas. Beserra et al. (2014) analisaram que a experiência se refere às informações que são obtidas sob a forma de dados, digitalizados por um operador humano ou através da interação da máquina com o seu ambiente. Logo, seguindo o contexto proposto por este autor, a aprendizagem de máquina pretende definir algoritmos e técnicas capazes de tornar uma máquina apta para tomar decisões, sem a intervenção humana, de forma eficiente.

A Gestão do Conhecimento (GC) é uma área interdisciplinar que propõe práticas e ferramentas para a captura, criação, armazenamento, disseminação, compartilhamento e uso do conhecimento. A indústria de software assume um papel importante dentro do setor de tecnologia da informação. A globalização dos mercados e o uso intensivo do conhecimento no desenvolvimento de novos produtos têm contribuído para que este setor se torne uma indústria de alta tecnologia, desejável para o crescimento e desenvolvimento econômico (BINUYO et al., 2015).

Assim, a principal matéria-prima de uma indústria de software é o conhecimento de seus membros. Eles se utilizam da criatividade e capacidade intelectual para o desenvolvimento de soluções adequadas para atender fins específicos (CHEN; RAGSDALL; O'BRIEN, 2014). Nesse cenário, notadamente, a GC é necessária devido ao fato do conhecimento, tanto dos desenvolvedores de software como de outros indivíduos envolvidos no projeto, estar diretamente relacionado ao desenvolvimento dos produtos, à gestão e à tecnologia (AURUM; DANESHGAR; WARD, 2008).

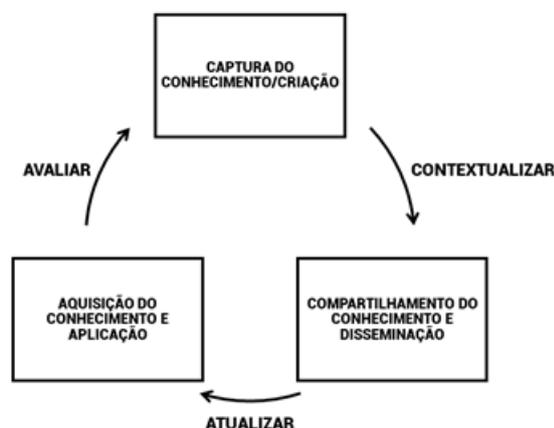
Tenório Jr. et. al. (2017) demonstram a importância do uso de ferramentas de GC no setor de tecnologia da informação. Os autores sugerem algumas vantagens que podem ser adquiridas com o uso de ferramentas de GC para esse tipo de indústria como, a melhoria da comunicação externa e o uso de conhecimento e de informações, contribuindo para o uso e aplicação do conhecimento tácito dos indivíduos. Para os autores o uso de ferramentas de GC irão possibilitar a existência dos estágios contidos em um ciclo de GC. Diante disso, se forma a seguinte questão de pesquisa: Como ocorre o ciclo de conhecimento nas ferramentas tecnológicas disponibilizadas pelo AM?

## 2 MATERIAIS E MÉTODOS

Foi utilizado metodologia de natureza aplicada através de pesquisa exploratória com revisão bibliográfica em Base de Dados Confiáveis (tais como, portal da CAPES, USP, Google Acadêmico entre outros) utilizando as palavras chave 'Ciclo do Conhecimento', 'Inteligência Artificial', 'Machine Learning' e 'Ferramentas Tecnológicas'. A ferramenta Google foi escolhida para a busca dos sites oficiais das ferramentas para a análise. Foram utilizadas as palavras-chaves "Machine Learning Tools" e "Machine Learning online tools" e selecionado apenas resultados que indicavam o site oficial de uma ferramenta. Foram selecionadas 13 ferramentas que mais condizem com esta pesquisa. Após a seleção as ferramentas foram separadas em categorias de tipos de tecnologia utilizada que se encontra na seção de resultados.

Monard et al. (2013) definem o AM como uma área de IA cujo objetivo é o desenvolvimento de técnicas computacionais sobre o aprendizado bem como a construção de sistemas capazes de adquirir conhecimento de forma automática. Já Mohri et al. (2012) define o AM como métodos computacionais usando a experiência para melhorar o desempenho ou fazer previsões precisas. O autor ainda complementa que o sucesso de um algoritmo de aprendizado depende dos dados usados, o AM é inerentemente relacionado à análise de dados e estatísticas. Em termos mais gerais, as técnicas de aprendizagem são métodos orientados por dados que combinam conceitos fundamentais em ciência da computação com ideias de estatísticas, probabilidade e otimização.

Para Dalkir (2005) o ciclo de informação do conhecimento pode ser visualizado como a rota que a informação segue para se transformar em um ativo estratégico valioso para a organização através de um ciclo de gestão do conhecimento. O autor ainda complementa que um dos principais processos de GC visa identificar e localizar fontes de conhecimento dentro da organização. Conhecimento valioso é então traduzido em forma explícita, muitas vezes referida como codificação do conhecimento, a fim de facilitar a disseminação mais ampla. Redes, práticas e incentivos são instituídos para facilitar a transferência de conhecimento de pessoa para pessoa, bem como conexões de conteúdo entre pessoa e conhecimento, a fim de resolver problemas, tomar decisões ou agir com base na melhor base de conhecimento possível. Uma vez que esse valioso conhecimento e know-how testados em campo é transferido para um repositório de conhecimento organizacional, diz-se que ele se torna parte da "memória corporativa". Isso às vezes também é chamado de "verdade fundamental".



**Figura 1** – Modelo traduzido do Ciclo do conhecimento  
Fonte: (DALKIR, 2005)

Dalkir (2005) descreve as principais etapas do ciclo do conhecimento. A captura de conhecimento refere-se à identificação e subsequente codificação de conhecimentos internos e know-how existentes



(normalmente despercebidos anteriormente) na organização e/ou conhecimento externo do ambiente. A criação de conhecimento é o desenvolvimento de novos conhecimentos e know-how - inovações que não tiveram existência prévia dentro da empresa. Quando o conhecimento é inventariado dessa maneira, o próximo passo crítico é apresentar uma avaliação em relação aos critérios de seleção que seguirão de perto os objetivos organizacionais.

Uma vez decidido que o conteúdo novo ou recém-identificado é de valor suficiente, o próximo passo é contextualizar esse conteúdo. Isso envolve a manutenção de um vínculo entre o conhecimento e os conhecedores desse conteúdo: o autor ou criador da ideia e os especialistas no assunto, bem como aqueles que acumularam experiência significativa no uso desse conteúdo. A contextualização também implica identificar os principais atributos do conteúdo para melhor corresponder a uma variedade de usuários. Por fim, a contextualização geralmente será bem-sucedida quando o novo conteúdo estiver firme, mas perfeitamente incorporado nos processos de negócios da organização. (DALKIR, 2005)

O ciclo de gerenciamento do conhecimento é então reiterado à medida que os usuários entendem e decidem usar o conteúdo. Os usuários validam a utilidade e sinalizam quando ficar desatualizada ou quando esse conhecimento não for aplicável. Os usuários ajudarão a validar o escopo do conteúdo ou a definir como as melhores práticas e lições aprendidas podem ser generalizadas. Eles também, muitas vezes, apresentam novos conteúdos, que podem contribuir para a próxima iteração do ciclo. (DALKIR, 2005)

### 3 RESULTADOS E DISCUSSÃO

Dentre os resultados coletados, foram selecionadas 13 ferramentas que utilizam algoritmos de AM para alcançar funcionalidades como predições, identificação visual e auditiva, leitura de texto (OCR), análise de dados, classificação e alguns outros. Abaixo temos um quadro com as ferramentas selecionadas:

**Quadro 1** - Lista de ferramentas de AM encontradas

| NOME [1]                                 | FUNÇÕES[2]   | TECNOLOGIAS [3] | ÁREAS DE ATUAÇÃO [4]  | CRIADOR [5]                   |
|--|--|-----------------|---|-------------------------------|
| <b>Scikit-Learn</b> <sup>1</sup>         | <i>Classification, Regression, Dimensionality reduction, Clustering</i>  | AM, Data Mining | Corporativo, Serviços Online, Pesquisa, Desenvolvimento, Estatística, Estudo, Aprendizado | Inria                         |
| <b>Shogun</b> <sup>2</sup>               | <i>Classification, Regression, Statistical Testing, Converter, Quantization, Clustering, Distance, Evaluation</i>              | AM              | Pesquisa, Desenvolvimento, Estatística, Estudo, Aprendizado                               | NumFOCUS                      |
| <b>Apache Mahout</b> <sup>3</sup>        | <i>Distributed Linear Algebra, Preprocessors, Regression, Clustering, Recommenders</i>   | AM              | Pesquisa, Estatística, Estudo   | Apache Software Foundation    |
| <b>Accord.net Framework</b> <sup>4</sup> | <i>Classification, Regression, Clustering, Distributions, Hypothesis Tests, Kernel Methods, Imaging, Audio, Signal, Vision</i> | AM              | Corporativo, Serviços Online, Pesquisa, Desenvolvimento, Estatística, Estudo, Aprendizado | Desenvolvedores Independentes |
| <b>Apache Spark ML lib</b> <sup>5</sup>  | <i>Classification, Regression, Decision trees, Recommendation, Clustering</i>  | AM              | Desenvolvimento, Estudo, Aprendizado  | Apache Software Foundation    |
| <b>H2O</b> <sup>6</sup>                  | <i>Regression, Classification, Prediction, Regularization, Quantization</i>  | AM              | Corporativo, Serviços Online, Pesquisa, Desenvolvimento, Estatística, Estudo, Aprendizado | H2O.ai                        |
| <b>Oryx 2</b> <sup>7</sup>               | <i>Recommendation, Classification, Regression, Clustering</i>  | AM              | Desenvolvimento, Estudo, Aprendizado  | Desenvolvedores Independentes |
| <b>GoLearn</b> <sup>8</sup>              | <i>Filtering, Classification, Regression</i>   | AM              | Desenvolvimento, Estudo, Aprendizado  | Desenvolvedores Independentes |
| <b>Weka 3</b> <sup>9</sup>               | <i>Filtering, Classification, Regression</i>   | AM, Data Mining | Pesquisa, Estatística, Estudo, Aprendizado  | University of Waikato         |



|                                    |  |                                       |   |        |
|------------------------------------|--|---------------------------------------|---|--------|
| <b>TensorFlow</b> <sup>10</sup>    | <i>Classification, Regression, Clustering, Distributions, Prediction, Kernel Methods, Imaging, Audio, Signal, Vision</i> | AM                                    | Corporativo, Serviços Online, Pesquisa, Desenvolvimento, Estatística, Estudo, Aprendizado | Google |
| <b>ConvNetJS</b> <sup>11</sup>     | <i>Classification, Regression, Imaging</i>   | <i>Deep Learning, Neural Networks</i> | Desenvolvimento, Estudo, Aprendizado  | MIT    |
| <b>ML Kit</b> <sup>12</sup>        | <i>Imaging, Audio, Vision, Text</i>  | AM                                    | Desenvolvimento, Aprendizado  | Google |
| <b>Cloud Auto ML</b> <sup>13</sup> | <i>Imaging, Audio, Vision, Text</i>  | AM                                    | Desenvolvimento, Aprendizado  | Google |

<sup>1</sup> "Scikit-learn." <http://scikit-learn.org/>. Acessado em 22 ago. 2018.

<sup>2</sup> "Shogun." <http://shogun-toolbox.org/>. Acessado em 22 ago. 2018.

<sup>3</sup> "Apache Mahout." <https://mahout.apache.org/>. Acessado em 22 ago. 2018.

<sup>4</sup> "Accord.NET." <http://accord-framework.net/>. Acessado em 22 ago. 2018.

<sup>5</sup> "MLlib." <https://spark.apache.org/mllib/>. Acessado em 22 ago. 2018.

<sup>6</sup> "H2O.ai." <https://www.h2o.ai/>. Acessado em 22 ago. 2018.

<sup>7</sup> "Oryx." <http://oryx.io/>. Acessado em 22 ago. 2018.

<sup>8</sup> "GoLearn." <https://github.com/sjwhitworth/golearn>. Acessado em 22 ago. 2018.

<sup>9</sup> "Weka." <https://www.cs.waikato.ac.nz/ml/weka/>. Acessado em 22 ago. 2018.

<sup>10</sup> "TensorFlow." <https://www.tensorflow.org/>. Acessado em 22 ago. 2018.

<sup>11</sup> "ConvNetJS" <https://cs.stanford.edu/people/karpathy/convnetjs/>. Acessado em 22 ago. 2018.

<sup>12</sup> "ML Kit | Google Developers." <https://developers.google.com/ml-kit/>. Acessado em 22 ago. 2018.

<sup>13</sup> "Cloud AutoML." <https://cloud.google.com/automl/>. Acessado em 22 ago. 2018.

Podemos visualizar nas informações do quadro 1 o nome da ferramenta [1], suas funções [2] encontradas no site oficial da ferramenta, as tecnologias utilizadas para a sua implementação e criação [3], suas áreas de atuação[4] e por fim, a empresa ou os indivíduos responsáveis pela sua criação [4]. Na coluna de tecnologias, podemos observar não só uma grande presença do AM como também do *Data Mining* (DM), ambas áreas que trabalham com algoritmos de alta complexidade e grande fluxo de dados. Além dessas tecnologias, temos a presença do *Deep Learning* (DL) e das *Neural Networks* (NN) que demonstram o crescimento e a amplitude da influência do ML em outras áreas, subáreas e novas tecnologias.

As ferramentas foram também separadas em áreas de atuação [6] que definem os possíveis clientes ou usuários das ferramentas a partir do que elas podem fornecer. Estas são:

- **Corporativo:** Empresas, corporações, indústrias, governo, saúde;
- **Serviços Online:** Aplicações web, serviços de streaming, serviços ao usuário;
- **Pesquisa:** Universidades, faculdades, fundações, professores, acadêmicos;
- **Desenvolvimento:** Desenvolvedores de aplicativos/web, empresas de desenvolvimento;
- **Estatística:** Matemáticos, pesquisadores de estatística, empresas de meteorologia;
- **Estudo:** Alunos da área de TI interessados em aprender a linguagem da ferramenta;
- **Aprendizado:** Pessoas leigas interessadas em ter uma experiência em ML.

A partir destas informações, e considerando principalmente a classificação das ferramentas por área de atuação e também as funções de cada ferramenta, podemos encontrar as possíveis áreas onde as ferramentas podem ser aplicadas no GC e no ciclo do conhecimento.

Sendo as principais:

- **Imagem, Áudio, Visão, Texto (OCR), Sinal**, que podem ser utilizadas nas etapas de captura/criação do conhecimento;
- **Recomendação, Classificação, Agrupamento e Regressão**, que podem ser utilizadas na etapa de contextualização;



- **Distribuição, Quantização e Teste estatístico**, que podem ser utilizados na etapa de compartilhamento e disseminação do conhecimento;
- **Predição**, que pode ser utilizado na etapa de atualização
- **Testes de Hipótese, Métodos de Kernel** que podem ser utilizados na aquisição e aplicação do conhecimento
- **Avaliação**, que pode ser utilizado na etapa de avaliação;

#### 4 CONSIDERAÇÕES FINAIS

Concluiu-se que as principais ferramentas de ML disponíveis atualmente podem auxiliar a GC e o Ciclo do Conhecimento com sistemas de automação, predição, recomendação e muitos outros. Espera-se, com este resultado, colaborar e incentivar o desenvolvimento de novas ferramentas para a GC que utilizem tecnologias baseadas em ML para a melhoria do Ciclo do Conhecimento.

#### REFERÊNCIAS

AURUM, A.; DANESHGAR, F.; WARD, J. **Investigating Knowledge Management practices in software development organizations - an Australian experience**. Information and Software Technology, v. 50, n. 6, p. 511–533, 2008.

BESERRA, C. A. et al. **Aplicação de Técnicas de Aprendizagem de Máquina em Objetos de Aprendizagem baseado em Software: um Mapeamento Sistemático a partir das Publicações do SBIE**. CINTED-UFRGS, Porto Alegre, v. 12, n. 1, p. 1-10, jul. 2014.

BINUYO, G. O. et al. **Evaluation of the Factors influencing the Indigenous Software Products Development in Nigeria**. International Journal on Advances in ICT for Emerging Regions (ICTer), v. 7, n. 3, p. 1–8, 2015.

CHEN, Hui.; RAGSDELL, Gillian.; O'BRIEN, Ann. **Identification of Tacit Knowledge Associated with Experience: a Chinese Software Industry Study**. In: European Conference on Knowledge Management, p. 1147-1155, 2014.

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. **NATURE**, London, v. 521, p. 436, mai. 2015.

LIAO, Shu-hsien. Knowledge management technologies and applications - literature review from 1995 to 2002. **Expert Systems with Applications**, TaiPei, Taiwan, v. 25, p. 155-164, jan. 2003.

MOHRI, Mehryar; ROSTAMIZADEH, Afshin; TALWALKAR, Ameet. **Foundations of machine learning**. 1 ed. Cambridge, Massachusetts: The MIT Press, 2012. 412 p.

NATH, M. D. et al. Extracting Meta Knowledge from Machine Learning and Data Mining Algorithms. **UIU Institutional Repository**, [S.L.], mai. 2018. Disponível em: <<http://dspace.uiu.ac.bd/handle/52243/222>>. Acesso em: 19 jul. 2018.

NIKOLOPOULOS, Chris. **Expert Systems: Introduction to First and Second Generation and Hybrid Knowledge Based Systems**. New York: Marcel Dekker, 1997. 331p.

RUSSELL, Stuart; NORVIG, Peter. **Artificial intelligence: a modern approach**. 3 ed. New Jersey: Pearson Education, 2009. 1062 p.



SELLITTO, Miguel Afonso. **INTELIGÊNCIA ARTIFICIAL: UMA APLICAÇÃO EM UMA INDÚSTRIA DE PROCESSO CONTÍNUO**. Gestão & Produção, São Leopoldo, RS, v. 9, n. 3, p.363-376, dez. 2002. Trimestral.

SILVA, A. M. L. et al. Descoberta de conhecimento através se métodos de aprendizagem de máquina supervisionados aplicados ao SIGAA/UFPI. **Revista de Sistemas e Computação**, Salvador, v. 7, n. 1, p. 68-78, jan./jun. 2007.

URBANO, G. et al. GESTÃO DO CONHECIMENTO NA INDÚSTRIA DE SOFTWARE: UM INSTRUMENTO PARA O LEVANTAMENTO DE REQUISITOS DE FERRAMENTAS. **VIII Mostra Interna de Trabalhos de Iniciação Científica**, Maringá, p. 2, out. 2016.