

**PROGRAMA INSTITUCIONAL DE BOLSAS DE INICIAÇÃO  
CIENTÍFICA E TECNOLÓGICA**

**UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY  
RIBEIRO**  
*Centro CCT*  
*Labotatório LCMAT*

**Plano de Trabalho para Renovação de Bolsa de Iniciação Científica**

**Bolsista:** Daniel Brito dos Santos

**Matricula:** 00119110393

**Orientadora:** Prof. Dra. Annabell Del Real Tamariz

**Curso:** Bacharelado em Ciência da Computação

**Título do Projeto:** Project-driven Data Science: Aprendendo e Mapeando

**Título do Plano de Trabalho:** Primeiros Passos na Inferência Ativa

**Fonte financiadora:** PIBIC/UENF

# 1 Justificativa

A ciência de dados é o campo que se ocupa do uso e desenvolvimento de técnicas computacionais e estatísticas para extrair ideias e conhecimento de dados de modo a auxiliar a tomada de decisões. Esse processo pode ser estruturado em atividades que podem ser agrupadas em um modelo de trabalho (Data Science Work-Model - DSWM) com quatro grande etapas: Preparação, Análise, *Deployment* e Comunicação [CFG<sup>+</sup>21].

No segundo ano desta pesquisa, nós desenvolvemos um projeto prático de dados, intitulado “AcaDem” seguindo essa estrutura. Construímos um software para extrair e processar os dados dos extratos acadêmicos dos alunos da UENF para construir um banco de dados local e calcular quantos alunos demandam cada disciplina em cada período letivo. Conseguimos, durante esse ano de pesquisa, desenvolver a maior parte das atividades do DSWM.

Entretanto, devido a sensibilidade dos dados contidos nos extratos, nós utilizamos o software apenas em um conjunto de quatro extratos voluntariamente fornecidos por colegas do curso de Computação da própria UENF. Desse modo não foi possível desenvolver a etapa de análise nesses dados. Atividade que substituímos pela análise de diversos outros *datasets* na preparação do um material didático, intitulado “Introdução a Ciência de Dados” para ensinar os primeiros passos que já tínhamos aprendido numa turma do curso de computação.

Também estudamos teoricamente os processos de análise, que segundo [WG16] é composto por transformação, visualização e modelagem de dados. Dessas três atividades conseguimos cobrir as duas primeiras nas atividades do segundo ano, mas a modelagem de dados se mostrou uma área de maior complexidade visto que ela em si é uma disciplina autônoma chamada de aprendizado de máquina.

Ou seja, enquanto a ciência de dados é campo que se concentra em coletar, gerenciar e analisar dados para extrair informações, o aprendizado de máquina (ML - Machine Learning) se ocupa dos algoritmos e técnicas que permitem às máquinas aprenderem com dados. Portanto, ML é uma vasta caixa de ferramentas utilizada pela ciência de dados tanto para gerar conhecimento, auxiliar em processamento de dados até na criação de sistemas que tomam decisões autônomas.

Considerando a diversidade da área de ML, entendemos ser produtivo concentrar este terceiro ano de pesquisa em um conceito influente e representativo dentro da área de ML, modelagem de dados e modelos de inferência. Um conceito que permita construir um panorama dos principais conceitos da área.

Em sua pesquisa inicial o bolsista encontrou e ficou extremamente interessado em explorar o campo de estudo da “Inferência Ativa”, ele acredita que este é o tema que une os seus interesses e parece estar no momento de pesquisa ideal para receber contribuição de novos pesquisadores. De modo que o bolsista pretende dedicar os próximos passos de sua carreira acadêmica nesse campo.

O conceito de Inferência Ativa busca formalizar o comportamento de sistemas dinâmicos, principalmente sistemas que apresentam de ação, percepção e aprendizado. A ideia central é que qualquer sistema dinâmico não estacionário para permanecer existindo necessita ter como ímpeto principal a minimização de sua energia livre. Desse modo, necessariamente o sistema necessita de algum tipo de representação do seu externo. O conceito de Inferência Ativa afirma que modelos de inferência são suficientes para modelar esse princípio e portanto modelar comportamento de sistemas muito mais complexos como braços robóticos e cérebros. [FSM<sup>+</sup>12, FRO<sup>+</sup>15, FFR<sup>+</sup>17]

Segundo [HMD<sup>+</sup>22], a pesquisa acadêmica sobre as aplicações da Inferência Ativa está em franca expansão nos mais diversos campos, dentre eles podemos citar:

- Modelagem de comportamento humano e animal [AVB<sup>+</sup>21, HPGF21, PRHF20]
- Neurociência cognitiva, especialmente tomada de decisão sob incerteza [SFM<sup>+</sup>15, SSS<sup>+</sup>20, SKS<sup>+</sup>21]
- Modelos computacionais de psicopatologia [MDFD12, SKS<sup>+</sup>21]
- Control Theory, área da engenharia que se ocupa de algoritmos e controladores para sistemas dinâmicos, desde braços robóticos até telecomunicações e foguetes. [BPC<sup>+</sup>22, BB19, MTSB20]
- Aprendizado de máquina por reforço [FSMF20, MTSB20, SBPF21, TBSB20, TMSB20]
- Cognição social [AVB<sup>+</sup>21, TP21, WT21]

Mais recentemente [KKT<sup>+</sup>22] demonstrou que uma cultura de células corticais demonstraram aprendizado a partir de estímulos previsíveis e imprevisíveis exatamente como sugere a teoria. O

que reforça a ideia de que o princípio da inferência ativa pode ser uma chave importante para compreender a cognição como fenômeno, tanto biológico, como talvez computacional. E que ela pode possivelmente explicar e permitir a geração de sistemas capazes de aprendizado e tomada de decisões.

Considerando a velocidade de desenvolvimento e os resultados relevantes que essa área tem apresentado mesmo sendo tão recente, entendemos que pode ser uma grande oportunidade para novos pesquisadores contribuírem com a sua consolidação.

Assim, no segundo ano concluímos um projeto prático de ciência de dados, AcaDem, neste próximo ano propomos concentrar o escopo na pesquisa acadêmica teórica de uma possível ferramenta de ponta para a ciência de dados, o aprendizado de máquina e a inteligência artificial.

Nessa primeira incursão buscaremos efetuar uma revisão bibliográfica dos artigos mais relevantes dessa área tendo em vista construir um mapa compreensível da área, buscando especialmente listar as lacunas e oportunidades de pesquisa como perguntas não respondidas e áreas pouco exploradas.

## 2 Objetivos

O objetivo dessa pesquisa é compreender os fundamentos teóricos e possíveis lacunas na pesquisa em Inferência Ativa.

## 3 Metodologia

Para atingir nosso objetivo de compreender a literatura do tópico propomos uma revisão bibliográfica exploratória. Seguiremos o seguinte algoritmo para desempenhá-la:

1. Buscar o termo “active inference” nas plataformas: Google Scholar, Semantics Scholar e Scopus. Essas plataformas são buscadores de artigos científicos que oferecem diversas ferramentas para nos auxiliar nesse busca inicial como a possibilidade de ordenar por número de citação, selecionar autores, filtrar datas de publicação.
2. Selecionar os cinco artigos que julgarmos mais representativos na área, a partir da busca exploratória nas três plataformas. Essa seleção se dará de acordo com o número total de citações de cada um deles, aliado a sua interconexão com os outros, os autores mais relevantes e principalmente a abrangência de sua proposta.
3. Efetuar uma leitura sistemática de cada um dos artigos selecionados buscando extrair seus principais conceitos. Construir um glossário dos termos citados, e um resumo de cada artigo.
4. Comparar os artefatos extraídos de cada artigo para construir uma análise comparativa da área. Nosso enfoque será especialmente localizar e explicitar os seus principais conceitos.

## 4 Etapas

Propomos as seguintes etapas para efetuar a metodologia proposta:

- A. Busca exploratória dos artigos.
- B. Leitura sistemática de cada artigo
- C. Análise comparativa dos achados na leitura
- D. Consolidação da revisão bibliográfica
- E. Elaboração do relatório final

	1º	2º	3º	4º	5º	6º	7º	8º	9º	10º	11º	12º
A. Busca exploratória dos artigos												
B. Leitura sistemática de cada artigo												
C. Análise comparativa dos achados na leitura												
D. Consolidação da revisão bibliográfica												
E. Elaboração do relatório final												

Tabela 1: Atividades a serem desempenhadas em cada mês de pesquisa

## Referências

- [AVB<sup>+</sup>21] R. A. Adams, P. Vincent, D. Benrimoh, K. J. Friston, and T. Parr. Everything is connected: Inference and attractors in delusions. *Schizophrenia Research*, Jul 2021.
- [BB19] M. Baltieri and C. L. Buckley. Pid control as a process of active inference with linear generative models. *Entropy*, Mar 2019.
- [BPC<sup>+</sup>22] M. Baïoumy, C. Pezzato, C. H. Corbato, N. Hawes, and R. Ferrari. Towards stochastic fault-tolerant control using precision learning and active inference. In *Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, pages 681–691. Springer International Publishing, 2022.
- [CFGT21] Anamaria Crisan, Brittany Fiore-Gartland, and Melanie Tory. Passing the Data Baton : A Retrospective Analysis on Data Science Work and Workers. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1860–1870, February 2021.
- [FFR<sup>+</sup>17] Karl Friston, Thomas FitzGerald, Francesco Rigoli, Philipp Schwartenbeck, and Giovanni Pezzulo. Active Inference: A Process Theory. *Neural Computation*, 29(1):1–49, January 2017.
- [FRO<sup>+</sup>15] Karl Friston, Francesco Rigoli, Dimitri Ognibene, Christoph Mathys, Thomas Fitzgerald, and Giovanni Pezzulo. Active inference and epistemic value. *Cognitive Neuroscience*, 6(4):187–214, October 2015. Publisher: Routledge .eprint: <https://doi.org/10.1080/17588928.2015.1020053>.
- [FSM12] Karl Friston, Spyridon Samothrakis, and Read Montague. Active inference and agency: optimal control without cost functions. *Biological Cybernetics*, 106(8):523–541, October 2012.
- [FSMF20] Z. Fountas, N. Sajid, P. A. M. Mediano, and K. J. Friston. Deep active inference agents using monte-carlo methods. In *Advances in Neural Information Processing Systems*. Advances in Neural Information Processing Systems, 2020.
- [HMD<sup>+</sup>22] Conor Heins, Beren Millidge, Daphne Demekas, Brennan Klein, Karl Friston, Iain D. Couzin, and Alexander Tschantz. pymdp: A Python library for active inference in discrete state spaces. *Journal of Open Source Software*, 7(73):4098, May 2022.
- [HPGF21] E. Holmes, T. Parr, T. D. Griffiths, and K. J. Friston. Active inference, selective attention, and the cocktail party problem. *Neuroscience & Biobehavioral Reviews*, Sep 2021.
- [KKT<sup>+</sup>22] Brett J. Kagan, Andy C. Kitchen, Nhi T. Tran, Forough Habibollahi, Moein Khajehnejad, Bradyn J. Parker, Anjali Bhat, Ben Rollo, Adeel Razi, and Karl J. Friston. In vitro neurons learn and exhibit sentience when embodied in a simulated game-world. *Neuron*, Oct 2022.
- [MDFD12] P. R. Montague, R. J. Dolan, K. J. Friston, and P. Dayan. Computational psychiatry. *Trends in Cognitive Sciences*, Jan 2012.
- [MTSB20] B. Millidge, A. Tschantz, A. K. Seth, and C. L. Buckley. On the relationship between active inference and control as inference. In *International Workshop on Active Inference*, pages 3–11. Springer International Publishing, 2020.

- [PRHF20] T. Parr, R. V. Rikhye, M. M. Halassa, and K. J. Friston. Prefrontal computation as active inference. *Cerebral Cortex*, Mar 2020.
- [SBPF21] N. Sajid, P. J. Ball, T. Parr, and K. J. Friston. Active inference: Demystified and compared. *Neural Computation*, Mar 2021.
- [SFM<sup>+</sup>15] P. Schwartenbeck, T. FitzGerald, C. Mathys, R. Dolan, and K. J. Friston. The dopaminergic midbrain encodes the expected certainty about desired outcomes. *Cerebral Cortex*, Sep 2015.
- [SKS<sup>+</sup>21] R. Smith, N. Kirlic, J. L. Stewart, J. Touthang, R. Kuplicki, S. S. Khalsa, J. Feinstein, M. P. Paulus, and R. L. Aupperle. Greater decision uncertainty characterizes a transdiagnostic patient sample during approach-avoidance conflict: A computational modelling approach. *Journal of Psychiatry & Neuroscience*, Jan 2021.
- [SSS<sup>+</sup>20] R. Smith, P. Schwartenbeck, J. L. Stewart, R. Kuplicki, H. Ekhtiari, M. P. Paulus, and Tulsa 1000 Investigators. Imprecise action selection in substance use disorder: Evidence for active learning impairments when solving the explore-exploit dilemma. *Drug and Alcohol Dependence*, Dec 2020.
- [TBSB20] A. Tschantz, M. Baltieri, A. K. Seth, and C. L. Buckley. Scaling active inference. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2020.
- [TMSB20] A. Tschantz, B. Millidge, A. K. Seth, and C. L. Buckley. Reinforcement learning through active inference. In *Bridging AI and Cognitive Science at the International Conference on Learning Representations*. Bridging AI and Cognitive Science at the International Conference on Learning Representations, 2020.
- [TP21] R. Tison and P. Poirier. Communication as socially extended active inference: An ecological approach to communicative behavior. *Ecological Psychology*, Apr 2021.
- [WG16] Hadley Wickham and Garrett Grolemund. *R for Data Science*. O’Reilly Media, Inc., 2016.
- [WT21] N. Wirkuttis and J. Tani. Leading or following? dyadic robot imitative interaction using the active inference framework. *IEEE Robotics and Automation Letters*, Apr 2021.