

**PREDIÇÃO DE RISCO DE ATRASO NO DESENVOLVIMENTO
NEUROPSICOMOTOR NOS PRIMEIROS MIL DIAS DE VIDA UTILIZANDO
APRENDIZADO DE MÁQUINA: COORTE BRISA**

SÃO LUÍS, MA

2026

KATIA SUSANA AZEVEDO SILVA

**PREDIÇÃO DE RISCO DE ATRASO NO DESENVOLVIMENTO
NEUROPSICOMOTOR NOS PRIMEIROS MIL DIAS DE VIDA UTILIZANDO
APRENDIZADO DE MÁQUINA: COORTE BRISA**

Dissertação apresentada ao Programa de Pós-Graduação em Saúde Coletiva da Universidade Federal do Maranhão como requisito parcial para obtenção do título de Mestre em Saúde Coletiva.

Orientador: Prof. Dr. Bruno Feres de Souza
Coorientadora: Prof.^a Dr.^a Cecília Cláudia
Costa Ribeiro de Almeida

SÃO LUÍS, MA

2026

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Diretoria Integrada de Bibliotecas/UFMA

Azevedo Silva, Katia Susana.

Predição de Risco de Atraso no Desenvolvimento
Neuropsicomotor nos Primeiros Mil Dias De Vida Utilizando
Aprendizado de Máquina: Coorte BRISA / Katia Susana
Azevedo Silva. - 2026.

166 p.

Coorientador(a) 1: Cecilia Cláudia Costa Ribeiro.

Orientador(a): Bruno Feres de Souza.

Dissertação (Mestrado) - Programa de Pós-graduação em
Saúde Coletiva/ccbs, Universidade Federal do Maranhão, São
Luís - Ma, 2026.

1. Desenvolvimento Infantil. 2. Aprendizado de
Máquina. 3. Modelos Preditivos. 4. Saúde da Criança. I.
Costa Ribeiro, Cecilia Cláudia. II. Feres de Souza,
Bruno. III. Título.

**PREDIÇÃO DE RISCO DE ATRASO NO DESENVOLVIMENTO
NEUROPSICOMOTOR NOS PRIMEIROS MIL DIAS DE VIDA UTILIZANDO
APRENDIZADO DE MÁQUINA: COORTE BRISA**

Katia Susana Azevedo Silva

Dissertação de Mestrado aprovada em 27 de fevereiro de 2026 pela banca examinadora
constituída dos seguintes membros:

Banca Examinadora:

Prof. Dr. Bruno Feres de Souza
Orientador
Universidade Federal do Maranhão

Prof.^a Dr.^a Cecília Cláudia Costa Ribeiro
Coorientadora
Universidade Federal do Maranhão

Prof. Dr. João Otávio Bandeira Diniz
Examinador Externo
Instituto Federal do Maranhão

Prof.^a Dr.^a Poliana Cristina de Almeida Fonseca Viola
Examinadora Interna
Universidade Federal do Maranhão

AGRADECIMENTOS

A Deus, fonte de força, sabedoria, proteção e saúde, por ter me sustentado diariamente e permitido que eu permanecesse firme no meu propósito.

À minha mãe, Alzira Azevedo pelo amor incondicional, pelas orações diárias e pelo apoio constante em todos os momentos. Sua força e dedicação sempre foram meu maior exemplo.

Aos meus filhos, Vinicius Azevedo e Maria Victoria, razão da minha vida, por serem meu incentivo diário e minha maior motivação. Agradeço por todo amor, cuidado, compreensão e ajuda em todos os momentos, especialmente nos dias difíceis.

Aos meus irmãos e familiares pelo carinho, compreensão e incentivo ao longo desta data trajetória, especialmente nos períodos de ausência e dedicação aos estudos.

Ao meu orientador, Prof. Dr. Bruno Feres, pela competência, dedicação, paciência e incentivo constante na condução deste trabalho.

À minha Coorientadora, Prof.^a Dr.^a Cecilia Claudia, pela oportunidade, apoio e contribuições no trabalho.

Aos colegas de trabalho, pela compreensão, apoio e palavras de encorajamento, que tornaram a jornada mais leve e possível.

Aos colegas da turma de mestrado e doutorado, pela parceria, pelas trocas de conhecimento, pelo companheirismo e apoio mútuo durante todo o curso.

A todos os professores do PPGSC, pelos ensinamentos que contribuíram para minha formação acadêmica.

À banca examinadora pela disponibilidade, dedicação e contribuições na avaliação deste trabalho.

À Universidade Federal do Maranhão, pela oportunidade de formação no mestrado e pelo suporte institucional.

Ao Programa de Pós-Graduação em Saúde Coletiva, pelo acolhimento e pela excelência acadêmica ao longo desta trajetória.

À Secretaria do PPGSC, pelo apoio e orientações no decorrer do curso.

“Tudo posso naquele que me fortalece”.

Filipenses 4:13

LISTA DE FIGURAS

Figura 1 Fluxograma da Coorte Pré-Natal BRISA, São Luís – MA -----	32
Figura 2 Modelo teórico do estudo com validação interna -----	39
 Artigo	
Figura 1 Distribuição da densidade do poder preditivo univariado (AUC) das variáveis explicativas para o desfecho de risco de atraso no desenvolvimento neuropsicomotor -----	60
Figura 2 Importância das variáveis preditoras para o cenário de risco cognitivo de atraso no desenvolvimento neuropsicomotor aos dois anos -----	61
Figura 3 Importância das variáveis preditoras para o cenário de risco comunicação agrupados de atraso no desenvolvimento neuropsicomotor aos dois anos -----	62
Figura Suplementar 1 Importância das variáveis preditoras para o cenário de risco motor agrupados de atraso no desenvolvimento neuropsicomotor aos dois anos -----	63
Figura Suplementar 2 Importância das variáveis preditoras para o cenário de risco global agrupados de atraso no desenvolvimento neuropsicomotor aos dois anos -----	64

LISTA DE TABELAS

Tabela 1. Descrição das habilidades do desenvolvimento neuropsicomotor de acordo com a idade -----	18
Tabela 2 – Características dos estudos internacionais de Aprendizado de Máquina aplicados à predição de risco do atraso no desenvolvimento neuropsicomotor-----	27
Tabela 3 – Bloco de variáveis selecionadas para predição do desenvolvimento neuropsicomotor-----	36
Tabela 4 – Modelo de Matriz de Confusão par algoritmo de classificação binária-----	42

Artigo

Tabela 1 Métricas de desempenho dos algoritmos de aprendizado de máquina para classificação do risco cognitivo; comunicação expressiva, receptiva e geral; motor fino, grosso e motor geral; e risco global na coorte BRISA, São Luís - MA, Brasil, n = 972-----	58
Tabela 2 Comparação das métricas de desempenho preditivo entre o presente estudo e a literatura -----	65

LISTA DE ABREVIATURAS E SIGLAS

AIMS – *Alberta Infant Motor Scale*

AM – Aprendizado de Máquina

APS – Atenção Primária à Saúde

AUC – Área sob a curva

BSID - III – *Bayley Scales of Infant and Toddler Development – Third Edition*

BPN – Baixo peso ao nascer

BRISA – Estudo de Coortes Brasileiras de Nascimento de Ribeirão Preto e São Luís

CEB – Classificação Econômica do Brasil

CES-D - Escala de Rastreamento Populacional para Depressão do Centro de Estudos Epidemiológicos

DNPM – Desenvolvimento Neuropsicomotor

DOHaD – *Developmental Origins of Health and Disease*

IMC – Índice de Massa Corporal

IPAQ – Questionário internacional de atividade física

MOS – Escala de Rede de Apoio do *Medical Outcomes Study*

MS – Ministério da Saúde

ODS – Objetivos de Desenvolvimento Sustentável

OMS – Organização Mundial da Saúde

PNAISC - Política Nacional de Atenção Integral à Saúde Criança

PSS14 – Escala de Estresse Percebido

RF – Random Forest

RN - Recém-nascido

RNA – Redes Neurais Artificiais

SHAP - Shapley Additive Explanations

SUS – Sistema Único de Saúde

SVM – Suporte de Vetor de Máquina

TEA – Transtorno do Espectro Autista

TCLE – Termo de Consentimento Livre e Esclarecido

TDAH – Transtorno do Déficit de Atenção com Hiperatividade

SILVA, Katia Susana Azevedo Silva. **Predição de Risco de Atraso no Desenvolvimento Neuropsicomotor nos Primeiros Mil Dias de Vida Utilizando Aprendizado de Máquina: Coorte BRISA**, 2026, Dissertação (Mestrado em Saúde Coletiva) – Programa de Pós-graduação em Saúde Coletiva, Universidade Federal do Maranhão, São Luís, 166p

RESUMO

Introdução: Os primeiros mil dias de vida representam uma janela de plasticidade cerebral crítica, na qual o desenvolvimento neuropsicomotor (DNPM) é sensível a influências biológicas e psicossociais. A predição precoce de riscos nesse período é fundamental para intervenções oportunas, embora constitua um desafio complexo em populações de base comunitária. **Objetivo:** Desenvolver e avaliar modelos de Aprendizado de Máquina (AM) para prever o risco de atraso no DNPM em crianças nos primeiros mil dias de vida. **Métodos:** Estudo prospectivo com dados de 972 binômios mãe-criança da coorte BRISA (São Luís – MA). O desfecho foi o risco de atraso avaliado nos domínios cognitivo, comunicação e motor avaliados pela escala Bayley III. Foram testados os algoritmos Regressão Logística Penalizada (PLR), *Random Forest* (RF), *Support Vector Machines* (SVM) com kernel radial e Redes Neurais Artificiais (RNA), e o classificador nulo. O pré-processamento incluiu a imputação de dados faltantes, codificação de variáveis categóricas por *one-shot encoding* e padronização de variáveis numéricas. Para a avaliação dos modelos de AM gerados, foi utilizada validação cruzada *K-fold* ($k = 5$), com 10 repetições. Como métricas de desempenho utilizaram-se a área sob curva AUC-ROC (AUC), a sensibilidade, a especificidade, a acurácia, a precisão e o F1-score. A interpretabilidade dos modelos foi analisada pela técnica *SHapley Additive Explanations* (SHAP). **Resultados:** A maior prevalência de risco de atraso ocorreu na comunicação expressiva (44,1%). Os modelos apresentaram desempenho moderado com AUC variando entre 0,533 e 0,603, sendo a PLR o algoritmo mais estável, atingindo AUC de 0,603 no domínio motor fino. A análise SHAP identificou como principais preditores: história familiar de epilepsia, sexo da criança, sintomas depressivos maternos, violência psicológica e suporte social, escolaridade materna e renda familiar. **Conclusão:** Modelos de AM demonstraram capacidade moderada de discriminação em amostras populacionais heterogêneas, evidenciando que o aumento da complexidade algorítmica não superou a regressão logística. Os achados reforçam a natureza multifatorial do desenvolvimento, destacando o impacto de determinantes psicossociais e biológicos, contribuindo para o desenvolvimento de ferramentas de triagem aplicáveis à Atenção Primária à Saúde.

Palavras-chave: Desenvolvimento Infantil; Aprendizado de Máquina; Modelos Preditivos; Saúde da Criança.

SILVA, Katia Susana Azevedo Silva. **Predição de Risco de Atraso no Desenvolvimento Neuropsicomotor nos Primeiros Mil Dias De Vida Utilizando Aprendizado de Máquina: Coorte BRISA**, 2026, Dissertação (Mestrado em Saúde Coletiva) – Programa de Pós-graduação em Saúde Coletiva, Universidade Federal do Maranhão, São Luís, 166p

ABSTRACT

Introduction: The first thousand days of life represent a critical window of brain plasticity, in which neuropsychomotor development (NPMD) is sensitive to biological and psychosocial influences. Early prediction of risks during this period is fundamental for timely interventions, although it constitutes a complex challenge in community-based populations. **Objective:** To develop and evaluate Machine Learning (ML) models to predict the risk of NPMD delay in children during the first thousand days of life. **Methods:** Prospective study with data from 972 mother-child pairs from the BRISA cohort (São Luís – MA). The outcome was the risk of delay assessed in the cognitive, communication and motor domains evaluated by the Bayley III scale. The algorithms Penalized Logistic Regression (PLR), Random Forest (RF), Support Vector Machines (SVM) with radial kernel and Artificial Neural Networks (ANN), and the null classifier were tested. Preprocessing included imputation of missing data, encoding of categorical variables by one-shot encoding and standardization of numerical variables. For the evaluation of the generated ML models, K-fold cross-validation ($k = 5$) with 10 repetitions was used. As performances metrics, the area under the AUC-ROC curve (AUC), sensitivity, specificity, accuracy, precision and F1-score were used. The interpretability of the models was analyzed using the SHapley Additive Explanations (SHAP) technique. **Results:** The highest prevalence of risk of delay occurred in expressive communication (44.1%). The models showed moderate performance with AUC ranging between 0.533 and 0.603, with PLR being the most stable algorithm, reaching an AUC of 0.603 in the fine motor domain. SHAP analysis identified as main predictors: family history of epilepsy, child's sex, maternal depressive symptoms, psychological violence and social support, maternal education and family income. **Conclusion:** ML models demonstrated moderate discrimination ability in heterogeneous population samples, showing that increased algorithmic complexity did not surpass logistic regression. The findings reinforce the multifactorial nature of development, highlighting the impact of psychosocial and biological determinants, contributing to the development of screening tools applicable to Primary Health Care.

Keywords: Child development; Machine Learning; Predictive models; Child health

SUMÁRIO

1. INTRODUÇÃO	12
2. OBJETIVOS	15
2.1 Objetivo Geral.....	15
2.2 Objetivos Específicos	15
3. REFERENCIAL TEÓRICO	16
3.1 Desenvolvimento Neuropsicomotor nos Primeiros Mil Dias de Vida	16
3.2 Avaliação do Desenvolvimento Neuropsicomotor e a Escala Bayley III	19
3.3 Atraso no Desenvolvimento Neuropsicomotor.....	20
3.4 Preditores do risco de Atraso do Desenvolvimento Neuropsicomotor.....	21
3.5 Aprendizado de Máquina na Saúde	23
4. MÉTODO	30
4.1 Desenho do estudo	30
4.2 Local de Estudo.....	30
4.3 População	31
4.5 Critérios de inclusão e exclusão	31
4.5 Procedimentos de Coleta de Dados.....	32
4.6 Variáveis do Estudo	35
4.7 Processamento e Análise de Dados	38
4.8 Aspectos Éticos e Legais	43
5 RESULTADOS	44
5.1 Artigo	44
6 CONSIDERAÇÕES FINAIS.....	68
REFERÊNCIAS	69
ANEXOS	74