

Estudo da mortalidade infantil e sua relação com saneamento básico no Brasil

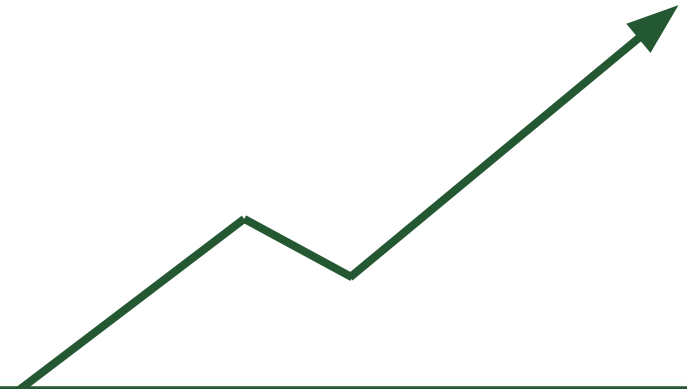
Leonardo Silva Pinto
Maria Paula Henriques Prandt



 @1workshop_ia.unifesp

 wia.ict.unifesp.br

- A TMI é um dos principais índices socioeconômicos;
- Tem-se como motivação corroborar com estudos já existentes sobre o assunto;
- Um dos principais desafios foi lidar com dados faltantes (NaN) e comprovar algumas hipóteses de relações entre fatores;
- Na literatura, são feitos estudos com cidades e regiões mais específicas a fim de garantir maior veracidade e acurácia das informações;



Encontrar parâmetros socioeconômicos que se relacionam com a Taxa de Mortalidade Infantil utilizando Inteligência Artificial (IA)

A base de dados utilizada é o Registro Administrativo Total de 2012 a 2017 e se encontra no acervo do site Atlas Brasil

Segmentação 1

SNIS_PAGUA (Percentual da população urbana residentes em domicílios ligados a rede de abastecimento de água)

PINTERDRSAI (Percentual de internações por doenças relacionadas ao saneamento ambiental inadequado)

SNIS_PESGOTO (Percentual da população urbana residente em domicílios ligados à rede de esgotamento sanitário)

SNIS_PESGTRA (Percentual de esgoto tratado)

PDEFSAN (Percentual de pessoas inscritas no Cadastro Único sem abastecimento de água, esgotamento sanitário e coleta de lixo adequadas)

PDEFAGUA (Percentual de pessoas inscritas no Cadastro Único sem abastecimento de água adequado)

Segmentação 2

TTREVA_EF_TOTAL (Taxa de evasão no ensino fundamental)

TTREVA_EM_TOTAL (Taxa de evasão no ensino médio)

DOCSUP_EF_TOTAL (Porcentagem de docentes do ensino fundamental com formação adequada)

DOCSUP_EM_TOTAL (Porcentagem de docentes do ensino médio com formação adequada).

Segmentação 3

REN_PIBPC_D (Produto interno bruto per capita)

REN_TRPCBF_D (Transferência per capita do Bolsa Família)

PIND_POS (Percentual de extremamente pobres no Cadastro Único pós Bolsa Família).

Segmentação Total

PINTERDRSAI: (Percentual de internações por doenças relacionadas ao saneamento ambiental inadequado)

DOCSUP_EF_TOTAL: (Porcentagem de docentes do ensino fundamental com formação adequada)

DOCSUP_EM_TOTAL: (Porcentagem de docentes do ensino fundamental com formação adequada)

TXNASC7C: (Percentual de nascidos vivos com pelo menos sete consultas de pré-natal)

TXNBAIXOP: (Percentual de nascidos vivos com baixo peso ao nascer)

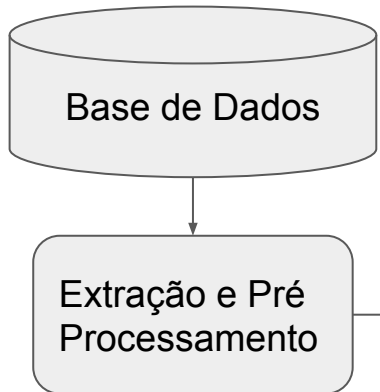
TXMAE10A14: (Percentual de meninas de 10 a 14 anos de idade que tiveram filhos)

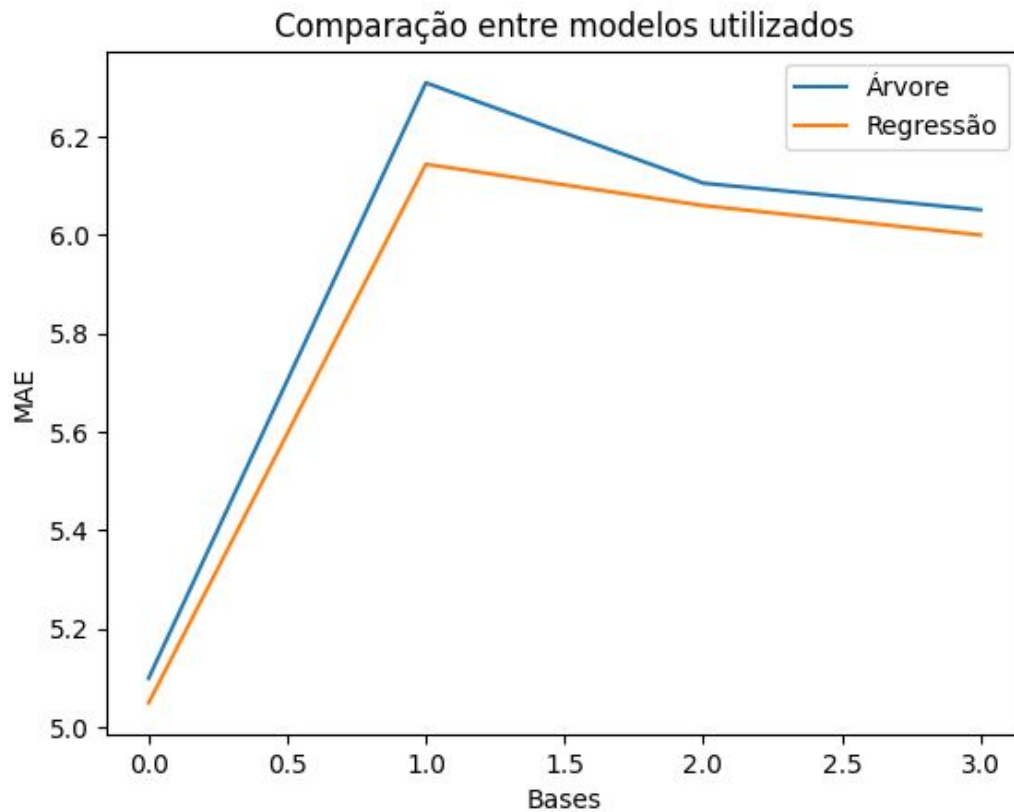
TXMAE15A19: (Percentual de adolescentes de 15 a 17 anos de idade que tiveram filhos)

REN_PIBPC_D (Produto interno bruto per capita)

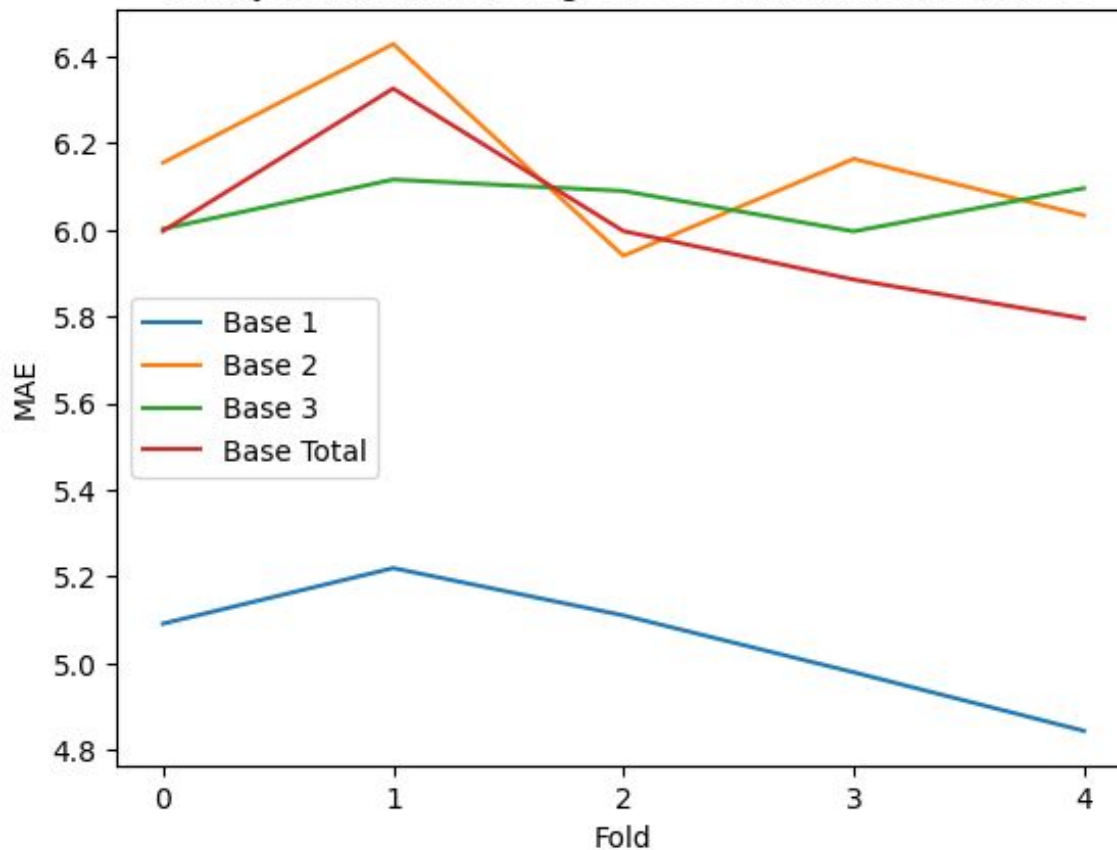
REN_TRPCBF_D (Transferência per capita do Bolsa Família)

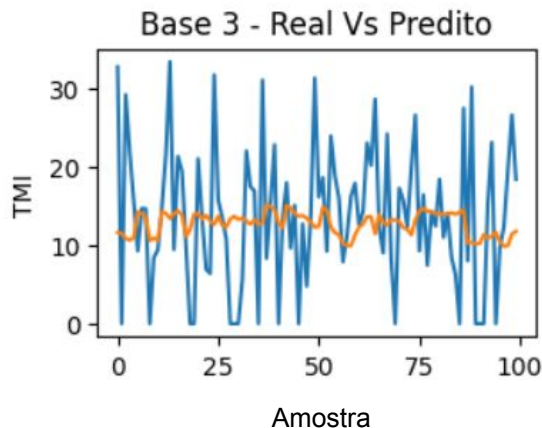
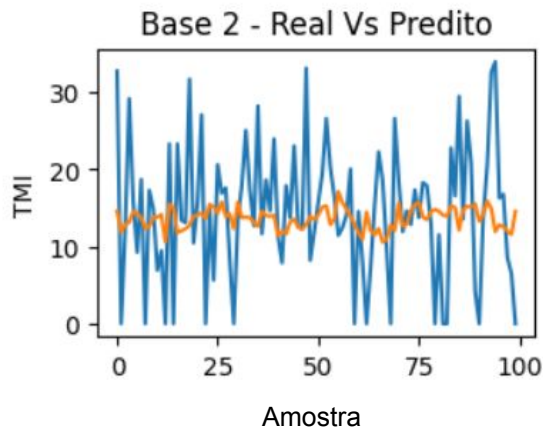
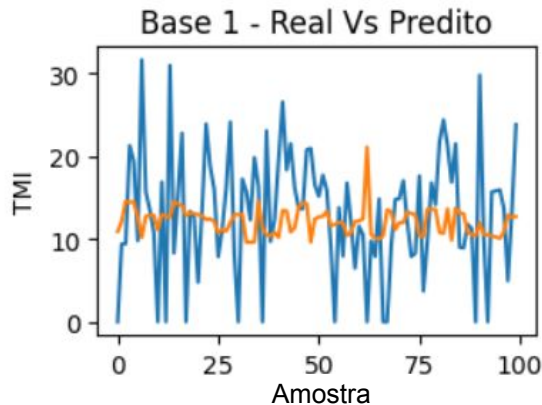
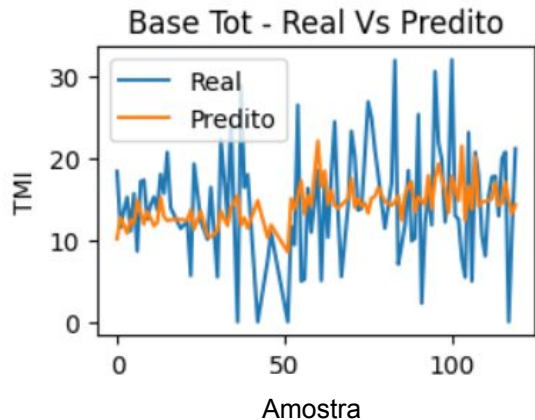
- A técnica usada foi regressão por meio de árvore de decisão e RNAs (Rede Neurais Artificiais)
- O protocolo de validação utilizado para a regressão é o de validação cruzada K-fold
 - **MAE - Erro Médio Absoluto**



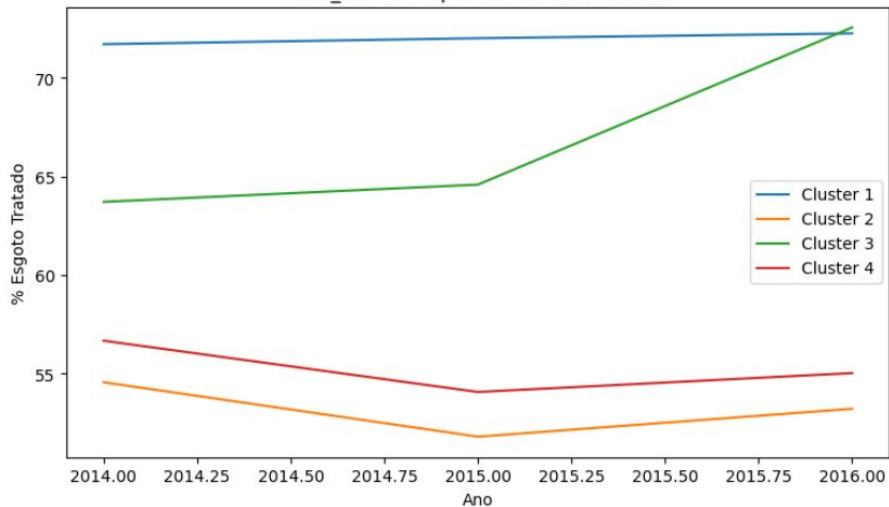


Variação da MAE nas regressões com diferentes bases

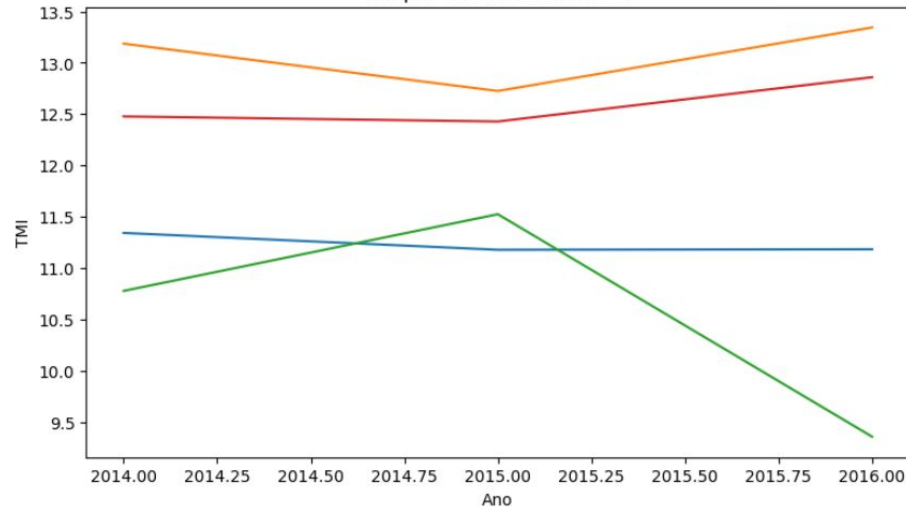




SNIS_PESGOTO por ano de cada cluster



TMI por ano de cada cluster



- Dados da classe “sanitários” foram os mais satisfatórios numa análise comparativa;
- Ainda assim, a regressão segue com resultados com uma MAE grande que pode se dar pelo viés de alguns dados;
- Para trabalhos futuros, cabe a busca por métodos para tratamento de dados NaN com grande desvio padrão;

- Professor Dr. Fabio Faria
- Professora Dr. Juliana Souza Scriptore