

**FUNDAÇÃO GETULIO VARGAS – ESCOLA DE MATEMÁTICA APLICADA  
INTELIGÊNCIA ARTIFICIAL E CIÊNCIA DE DADOS**

**BRUNO OTTONI ELOY VAZ  
MAURÍCIO GOUVÊA SILVA  
ROSSINI CORREIA LIMA NETO**

**O EMPREGO DA INTELIGÊNCIA ARTIFICIAL EM PROVEITO DO SISTEMA DE DEFESA  
NUCLEAR, BIOLÓGICA, QUÍMICA E RADIOLÓGICA DA MARINHA DO BRASIL  
(SISDEFNBQR-MB)**

**RIO DE JANEIRO  
2024**

## 1. INTRODUÇÃO

O SisDefNBQR-MB (Sistema de Defesa Nuclear, Biológica, Química e Radiológica da Marinha do Brasil) é o sistema da Marinha do Brasil responsável por coordenar e implementar ações de defesa e resposta a emergências nucleares, biológicas, químicas e radiológicas. O sistema integra diversos níveis e setores da Marinha para melhorar a capacidade de resposta rápida e proteção de instalações estratégicas.

Este trabalho se chama “Auxílio da Inteligência Artificial (IA) na detecção de ameaças do Sistema de Defesa NBQR” e tem o objetivo de criar uma inteligência artificial que, a partir de dados de localização e movimentação de pessoas, responderá informando se estar ocorrendo uma emergência ou não. A referida inteligência artificial pode ser particularmente importante para o SisDefNBQR-MB, visto que busca ajudar na detecção de situações que se assemelhem a emergências de natureza nuclear, biológica, química ou radiológica.

### 1.1 Contextualização

De acordo com Patriota (2022), as ameaças nucleares, biológicas, químicas e radiológicas estão presentes em nossas vidas e podem nos atingir quando menos esperamos. Assim, é importante conhecer tais ameaças e possuir meios para se proteger de seus efeitos nocivos.

Para atender a esta necessidade a Marinha do Brasil (MB) desenvolveu sua estrutura de Defesa Nuclear (N), Biológica (B), Química (Q) e Radiológica (R), hoje materializada em seu Sistema de Defesa NBQR (SisDefNBQR-MB). Este sistema foi estruturado para enfrentamento dos efeitos adversos advindos de qualquer uma das quatro vertentes N, B, Q ou R, simultaneamente ou não. Este conceito levou a uma tendência mundial de que as forças de resposta pudessem identificar a defesa contra os agentes N, B, Q e R por meio de uma sigla que unifica os conceitos de defesa nuclear, biológica, química e radiológica. Nesse contexto a MB desenvolveu estruturas especializadas e programas, como o Programa Nuclear da Marinha (PNM) e o Centro Tecnológico da Marinha em São Paulo (CTMSP). A evolução tecnológica e a necessidade de enfrentar novas ameaças levaram à criação de unidades

especializadas, como o Batalhão de Defesa NBQR. O Comando-Geral do Corpo de Fuzileiros Navais está inserido nesta estrutura.

A Brigada Real da Marinha foi a origem do Corpo de Fuzileiros Navais do Brasil (Marinha do Brasil, 2024). Criada em Portugal em 28 de agosto de 1797, por Alvará da rainha D. Maria I, chegou ao Rio de Janeiro em 7 de março de 1808, acompanhando a família real portuguesa que transmigrava para o Brasil.

Atualmente o Corpo de Fuzileiros Navais (CFN) está inserido na organização da Marinha, por intermédio do Comando-Geral do CFN (CGCFN) e de seu componente operativo, o Comando da Força de Fuzileiros da Esquadra (ComFFE), este último subordinado ao Comando de Operações Navais. O CFN atua na organização e preparo de Grupamentos Operativos de Fuzileiros Navais para realizar Operações Anfíbias de caráter de guerra naval, como as Operações Anfíbias e Ribeirinhas, passando pelas atividades de emprego limitado da força, como as Operações de Paz e as Garantias da Lei e da Ordem (GLO), até as atividades benignas, como as de apoio à defesa civil.

A permanente condição de pronto emprego dos Fuzileiros Navais atende à Estratégia Nacional de Defesa e demanda que as tropas da Força de Fuzileiros da Esquadra sigam um rigoroso ciclo de adestramento, que as capacita para atuar em todo o espectro das operações militares, desde um combate real, em operações de guerra naval, perpassando o emprego limitado da força, como nas operações de Garantia da Lei e da Ordem, e em operações benignas, como as operações de assistência humanitária. Assim, ameaças nucleares, biológicas, químicas e radiológicas enquadram-se no contexto de atuação do CFN.

## **1.2 Relevância do Tema**

De acordo com Andrade, E. Fisher, B. & Blunsden (2006), nos últimos anos técnicas de visão computacional e aprendizado de máquina (*machine learning*) têm sido aplicadas à modelagem e reconhecimento de atividades e interações humanas. Os domínios de aplicação dessas técnicas geralmente envolvem ambientes mais simples, de forma que o reconhecimento de atividades é focado na modelagem das ações e

interações de pequenos grupos de pessoas/objetos. No entanto, existem algumas tentativas de modelar grupos maiores de pessoas, multidões, que são principalmente baseadas em classificadores discriminativos. A análise dos movimentos e comportamentos de multidões é de particular interesse no domínio da vigilância e alerta. Em cenários onde centenas de câmeras são monitoradas por poucos operadores, a análise comportamental de multidões é útil como uma ferramenta de pré-triagem de vídeo.

Assim, para analisar uma multidão, um modelo deve lidar com uma grande variação de densidades e movimentos presentes em uma multidão real. Isso requer uma enorme quantidade de dados para permitir um bom aprendizado supervisionado/não supervisionado para modelos de multidão discriminativos ou generativos. Neste ambiente o que se vê são poucos ou nenhum exemplo de eventos de emergência/anormais a serem detectados. Assim, pode-se pensar em classificar os eventos como comportamentos normais ou anormais.

Desenvolver ferramentas automáticas que possam identificar "pontos problemáticos" em um local (a partir do conjunto completo de câmeras que o monitoram) e encaminhar as imagens correspondentes para uma sala de controle passa a ser um fator importante para diferentes áreas de negócio e, especialmente, a defesa. Tal abordagem deve permitir a gestão proativa de multidões, em vez do controle puramente reativo de incidentes, e ainda mantém o operador como o principal tomador de decisões no sistema.

Chaib (2015) informa que o SisDefNBQR-MB é responsável por responder a emergências NBQR nas instalações sensíveis, que incluem oficinas, usinas, laboratórios e protótipos relacionados à propulsão naval nuclear. Cada instalação possui seu próprio Plano de Emergência Local (PEL) que funciona como um Plano de Segurança Orgânica (PSO) específico para cada local.

O tempo de resposta é crucial em emergências NBQR, com equipes de pronta-resposta recomendadas para estarem presentes 24 horas. Assim, a defesa NBQR em instalações sensíveis é vital para a Marinha do Brasil, exigindo treinamento específico e coordenação eficiente para salvar vidas e mitigar os efeitos de emergências NBQR.

A contínua atualização e ausência de improvisação são essenciais para o sucesso dessas operações.

### **1.3 Escopo do Trabalho**

É neste sentido que surge o presente trabalho, que visa criar uma inteligência artificial que, a partir de dados de localização e movimentação de pessoas recebidos, responderá informando se tal situação é emergencial ou não. O programa será desenvolvido em Python e terá como base os estudos feitos nos módulos do curso “Inteligência Artificial e Ciência de Dados”.

Quanto ao emprego da IA na detecção de alguma ameaça NBQR, primeiramente foi identificado pela natureza da ameaça que a maior possibilidade de detecção seria em um agente químico (dado que os resultados sobre as vítimas de tal agente são de percepção imediata). Tal informação direcionou o foco nesse tipo de ameaça. Neste caso, seria considerada a dispersão das vítimas (que se evadiriam do local onde fora lançado o agente) e essa dispersão pode sinalizar uma possível dissipação de tal ataque, otimizando assim a detecção pelo Sistema de Defesa NBQR (SistDefNBQR).

## 2. REVISÃO DA LITERATURA

### 2.1 COMENTÁRIOS SOBRE AS FONTES DE INFORMAÇÃO

a) Módulos I, II e III do Curso

**Módulo I** - A primeira fonte a ser comentada é o próprio curso realizado até agora, em seus dois primeiros módulos. No primeiro deles, os alunos foram expostos aos modelos de distribuição estatísticos, tanto em provas matemáticas (iniciando por modelos de probabilidade, diagrama de Venn, propriedades de conjuntos, variáveis aleatórias, variáveis contínuas e discretas, esperança, variância, desvio padrão, covariância, dentre outras) de como chegar às suas fórmulas, como também suas aplicações (empíricas inclusive) para explicar o comportamento de aspectos naturais e comportamentos da sociedade em geral.

Tais demonstrações matemáticas contemplaram distribuições estatísticas como: Bernoulli, Binomial, Poisson, Geométrica, Uniforme, Exponencial, Normal e Normal Padrão. Materializando tais demonstrações, identificamos a constituição de fórmulas que demonstravam as esperanças, desvios padrões em cada uma dessas distribuições, bem como estimadores que pudessem se aproximar de tais valores (estimadores dentro de amostras ou conjunto de amostras dentro de uma população, que seguissem um comportamento possível de ser estudado por uma dessas distribuições). Assim, foi possível construir uma tabela no auxílio de identificar a melhor relação entre distribuição e comportamento, conforme a lista a seguir:

#### 1. Distribuição Bernoulli

- Notação:  $X \sim b(1, p)$
- Esperança Matemática:  $E(X) = 0 \cdot (1 - p) + 1 \cdot p = p$

#### 2. Distribuição Binomial

- Notação:  $X \sim B(n, p)$
- Esperança Matemática:  $E(X) = \sum_{i=0}^n \binom{n}{i} p^i (1 - p)^{n-i} = np$

### 3. Distribuição Poisson

- Notação:  $X \sim \text{Po}(\lambda)$
- Esperança Matemática:  $E(X) = \sum_{i=0}^{\infty} i \cdot \frac{e^{-\lambda} \lambda^i}{i!} = \lambda$

### 4. Distribuição Geométrica

- Notação:  $X \sim \text{Geometric}(p)$
- Esperança Matemática:  $E(X) = \sum_{i=1}^{\infty} i \cdot p(1-p)^{i-1} = \frac{1}{p}$

### 5. Distribuição Uniforme

- Notação:  $X \sim U(a, b)$
- Esperança Matemática:  $E(X) = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}$

### 6. Distribuição Exponencial

- Notação:  $X \sim \text{exp}(\lambda)$
- Esperança Matemática:  $E(X) = \int_0^{\infty} x \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$

### 7. Distribuição Normal

- Notação:  $X \sim N(\mu, \sigma^2)$
- Esperança Matemática:  $E(X) = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \mu$

### 8. Distribuição Normal Padrão

- Notação:  $X \sim N(0, 1)$
- Esperança Matemática:  $E(X) = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0$

A teoria, em boa parte, possibilitou que os alunos pudessem desenvolver a capacidade de melhor mensurar estimadores para esses parâmetros (esperança, desvio padrão e variância), de tal modo que fosse possível criar uma função (ou um conjunto delas) que pudesse chegar o mais próximo possível de explicar um comportamento específico de um evento. Alguns artifícios para se chegar a tal objetivo foram objeto de estudo, como o Teorema Central do Limite, a Máxima Verossimilhança entre Variáveis, distribuição T – Student, entre outros.

**Módulo II** - O segundo módulo focou mais em programação de forma mais prática. Os conhecimentos adquiridos no primeiro módulo foram aplicados à linguagem de programa R, uma linguagem de alto nível (um tipo de linguagem de programação que é mais voltada a linguagem humana que a linguagem de máquina). Para definir níveis de linguagem, entende-se o mais alto nível como o mais facilmente legível pelo homem e o mais baixo nível o de entendimento da máquina, basicamente o que é digitado de forma binária.

A grande vantagem do uso da linguagem R são suas funcionalidades para quem trabalha com estatística. A linguagem permite com poucas linhas de código ler uma grande quantidade de valores e torná-los informação útil ao usuário dessas informações. Além disso, práticas formas de criação de interações gráficas (tais como gráficos de múltiplos tipos, tabelas e muito mais) fazem dessa linguagem de amplo uso entre matemáticos e estatísticos.

Logo foi possível aplicar os conceitos aprendidos sobre distribuições estatísticas, suas médias, as variâncias e, também, sobre desvio padrão dentro desta parte do curso. Houve uma imersão em Python, uma das linguagens de programação mais utilizadas no mundo, capaz de interpretar dados de forma ágil, inclusive com a possibilidade de previsão de comportamentos.

Para consolidar o conhecimento dessa linguagem foi ministrado o seguinte roteiro:

1. Comandos Iniciais (Print, input) que consolidaram os fundamentos de programação;
2. Variáveis, tipos (Float, inteiro, string, booleano);
3. Operações aritméticas (soma, subtração, multiplicação, divisão, dentre outras);
4. Funções Padronização Dados (upper, lower, capitalize, dentre outras);
5. Funções Condicionais e Repetição (If, else, for, while); e
6. Definição de Funções (Pandas, NumPy, SciPy dentre outras).

Posteriormente, foi realizada uma aplicação desses conceitos, a fim de consolidar os conhecimentos de uma maneira prática. Tal aplicação foi através de modelos geoestatísticos, que, basicamente, tratam da criação de objetos como mapas com

eventos estatísticos sobre ele. Tal conhecimento possibilitou a criação de representações gráficas de ocorrências em dadas regiões (que pode auxiliar um decisor em uma tomada de decisão sobre algum aspecto específico, como por exemplo o posicionamento de viaturas policiais baseado em pontos de ocorrência de crimes em uma dada localidade), servindo de exemplificação para consolidar todos os conhecimentos até então adquiridos. Isto muito orientou a escolha do tema para este trabalho.

A Organização da crescente de conhecimento em Geoestatística seguiu três passos:

1. Fundamentos de Geoestatística;
2. Projetos Supervisionados: Aplicações em Cartografia; e
3. Projeto Supervisionado: Aplicações em Navegação.

O estudo serviu para exemplificar como a análise e a apresentação de dados poderia estar associada a aplicações militares, e seguiu alguns estudos de caso, quais sejam:

1. Estudo de Caso 1: Planejamento de Rotas;
2. Estudo de Caso 2: Manutenção Preventiva de Equipamentos;
3. Estudo de Caso 3: Reconhecimento de Atividades;
4. Estudo de Caso 4: Agrupamento [Clustering] de Objetos; e
5. Estudo de Caso 5: Uso de LLM [GPT, DALL-E, dentre outros].

**Módulo III** - O terceiro módulo do curso iniciou-se com foco na modelagem de dados. Nele estudou-se como seria construído um Data Warehouse (uma espécie de grande depósito de dados estruturados e organizados), como os dados seriam utilizados posteriormente e o uso deles em Dashboards para auxílio na tomada de decisões por um usuário final.

Assim, os alunos foram imersos no modus operandi de um conjunto de processos denominado ETL (Extract, Transform and Load), que consiste em extrair dados de fontes externas ao servidor da entidade, transformar o dado de forma a torná-lo estruturado e, posteriormente, carregá-lo dentro de um Data Warehouse. Na

sequência inicia-se o processo de construção de interface (Dashboard) para melhor apresentação dos dados para os tomadores de decisão.

Na consolidação do conhecimento desses processos, os alunos foram orientados a como modelar os dados, partindo do nível mais conceitual (Modelo Entidade Relacionamento), passando pelo Modelo Relacional (abrangendo o conceito de Modelagem Lógica) e chegando, finalmente, à Modelagem Dimensional, que dá ao usuário a capacidade de enxergar a estrutura de dados da melhor maneira para se construir um Data Warehouse (DW), com maior aplicabilidade.

A Linguagem SQL, rapidamente comentada, tem o papel de auxiliar os programadores a concretizarem os processos ETL, sendo voltada a realizar buscas dentro de bancos de dados e retornar apenas o essencial ao programador, tornando possível a criação estruturada do DW.

Outros conceitos foram abordados durante o curso, tais como uso do processo ELT (Extract Load and Transform) ao invés de um ETL, que seriam responsáveis por criar um Data Lake (seria um grande depósito de dados, com menor estruturação e organização que um DW, e que poderia abrigar uma quantidade maior de dados) ao invés de um Data Warehouse.

Neste contexto, entra a importância do Conhecimento em Power Business Intelligence (Power BI), outra disciplina ministrada, a qual possibilita o usuário construir da melhor maneira possível esse Dashboard para que auxilie o decisor com a melhor estruturação dos dados possível. A Business Intelligence possibilita insights para os gestores em níveis estratégicos. Suas ferramentas analisam dados passados e atuais e os apresentam num formato com visuais intuitivos e interativos.

Sendo assim, o curso destinou uma parte para o desenvolvimento da capacidade dos alunos de produzirem Dashboards interativos para auxiliar tomadas de decisões em suas Organizações Militares.

Concluído o estudo sobre o processo de modelagem dos dados (desde o processo ETL/ELT extraíndo dados de fontes externas, passando – os aos DW e depois

apresentando-os num Dashboard), na terceira semana o foco foi o desenvolvimento de conhecimentos a respeito de robótica. Os alunos puderam, empiricamente, realizar experimentos com uma placa de Arduino Uno, com seus sensores (equipamentos de Input) e seus dispositivos de Output. Neste contexto os alunos praticaram Python voltado a capturar dados do ambiente no qual o sensor da placa foi instalado, transmitindo os dados à placa (registrando-os), e, posteriormente, realizando uma análise por meio de uma interface visual que descrevia um comportamento do ambiente frente ao sensor.

Tais atividades relacionaram desde a captura sinais vitais humanos (ritmo cardíaco, saturação de oxigenação sanguínea) até condições meteorológicas (temperatura, distâncias de objetos ao sensor, campos magnéticos, sensor de movimento, dentre outros).

O desenvolvimento desta área serviu pra exemplificar aos alunos como a robótica poderia ser utilizada para capturar dados do ambiente, desde meteorologia até dados dentro de uma viatura militar ou mesmo de algum equipamento. Após a captura, um estudo relacionado poderia indicar um padrão de comportamento que pode vir a aumentar o desempenho desses equipamentos quanto a fatores externos ou do seu próprio funcionamento interno.

Um estudo massivo desses dados, sendo depois lançados em um grande banco de dados, poderia proporcionar um futuro emprego de Inteligência Artificial para melhor prever e assessorar possíveis tomadas de decisão sobre tais equipamentos e viaturas monitoradas, o que, por conseguinte, melhoraria as características desses bens militares, sendo possível, inclusive, conferir vantagem tática no campo de batalha. Um exemplo poderia ser a captura de quantidade de gases tóxicos emanados pelos veículos, seu controle e, assim, melhor uso de tais dispositivos pelos militares dessas viaturas, podendo elevar seu estado de conservação (bem como melhores condições de trabalho para os próprios militares) e uma possível vantagem tática.

## b) Publicações selecionadas

Na pesquisa realizada foram identificadas publicações que se julgaram importantes para o desenvolvimento deste trabalho. Entre as publicações identificadas destacam-se algumas, as quais estão comentadas a seguir.

**"A Review of Machine Learning and Deep Learning Techniques for Anomaly Detection in IoT Data"** - Apresenta uma revisão das técnicas de detecção de anomalias em dados de IoT, utilizando técnicas de aprendizado de máquina (*machine learning*) e aprendizado profundo (*deep learning*). No caso do *machine learning*, algoritmos como k-Nearest Neighbors (k-NN), Support Vector Machines (SVM), e árvores de decisão são comumente usados para detectar anomalias em fluxos de dados. Já no *deep learning*, as redes neurais profundas, como Autoencoders e Redes Neurais Convolucionais (CNNs), têm mostrado melhorias significativas na detecção de anomalias devido à sua capacidade de lidar com dados não estruturados e aprender representações complexas.

**"tegedet: An extensible Python Library for Anomaly Detection using Time-Evolving Graphs"** - Apresenta uma biblioteca Python para detecção de anomalias em abordagens de aprendizado não supervisionado, que se diferencia de outras por ser focada na extensibilidade e na integração de técnicas de detecção de anomalias baseadas em grafos dinâmicos. A biblioteca é desenhada para ser um ponto de referência em Python para a integração de técnicas de detecção de anomalias de última geração baseadas em grafos dinâmicos. A biblioteca proposta é, provavelmente, uma das poucas (ou única) disponível publicamente que pode ser estendida com outras técnicas de detecção de anomalias baseadas em grafos dinâmicos.

**"Detection of Emergency Events in Crowded Scenes"** - O artigo aborda técnicas de detecção de eventos anormais em multidões, um tópico crucial na vigilância de segurança. Propõe uma técnica para detectar eventos de emergência em multidões usando aprendizado não supervisionado para caracterizar o comportamento normal da multidão e identificar anomalias. Em termos metodológicos utiliza o fluxo óptico para observar o movimento da multidão e extração de características não

supervisionadas para codificar o comportamento normal da multidão. Para detecção de anomalias considera a segmentação do vídeo em blocos de frames para análise e a aplicação de análise de componentes principais (PCA) nos campos de fluxo óptico para reduzir a dimensionalidade. Trata-se de uma abordagem interessante na detecção de cenários de emergência observando a multidão em níveis locais e globais, que combina conceitos avançados de fluxo óptico, subespaços e HMMs para criar um método flexível capaz de lidar com uma ampla gama de densidades de pessoas, desde tráfego pedestre esparso até fluxos densos de multidões.

**"Anomaly Detection in Mixed High Dimensional Molecular Data"** – O artigo discute um sistema de detecção de anomalias baseado em aprendizado profundo aplicado a dados moleculares de alta dimensão. O sistema proposto utiliza redes neurais convolucionais (CNNs) para análise de vídeo em tempo real. Este sistema é escalável, rápido e flexível, podendo ser implementado em dispositivos como CCTV para detecção imediata de acidentes e desastres naturais. As características principais do sistema incluem: escalabilidade, rapidez e flexibilidade. O uso de *deep learning* mostrou-se uma abordagem interessante para melhorar a precisão e a velocidade na detecção de anomalias, oferecendo uma solução robusta para monitoramento de emergências e segurança.

**"Detection of Potentially Dangerous Situations involving Crowds using Image Processing"** – Este artigo aborda a detecção de situações potencialmente perigosas envolvendo multidões em locais de transporte público através do processamento de imagens. No caso, o objetivo está voltado à melhoria da segurança e a confiança do público que utiliza o transporte público, detectando automaticamente situações potencialmente perigosas usando processamento de imagens. A detecção tardia de eventos importantes, especialmente quando um único observador monitora muitas câmeras, foi o motivador deste estudo. A classificação usando textura e redes neurais funcionou adequadamente, ainda que necessite de melhorias em alguns aspectos.

**"Review on Computer Vision Techniques in Emergency Situations"** - O artigo discute o uso de técnicas de visão computacional em várias situações de emergência. Ele destaca a importância da consciência situacional e o uso de fotos e vídeos geolocalizados para entendimento e, a partir destas informações, tomar decisões

durante as emergências. O artigo aborda os desafios do *big data* e do excesso de informações, além da necessidade de ferramentas de visão computacional como suporte à decisão em emergências. A revisão é organizada em quatro eixos principais: tipos de emergências estudadas na visão computacional, objetivos que os algoritmos podem abordar, tipos de hardware necessários e algoritmos usados. Ele fornece uma visão geral do progresso da visão computacional em emergências, incluindo emergências naturais e provocadas pelo homem, e o papel da visão computacional em diferentes fases do gerenciamento de emergências, como prevenção, detecção, assistência e compreensão. O artigo também discute o uso de vários sensores visuais e sistemas de aquisição, e a aplicação de algoritmos de extração de características e classificação na análise de dados de emergência. Ele enfatiza a importância da escolha do sensor, da configuração de aquisição e dos algoritmos no desenho de sistemas eficazes de visão computacional para o gerenciamento de emergências.

## **2.2 RELAÇÃO DA LITERATURA COM O PROJETO**

Para criar uma inteligência artificial que, a partir de dados de localização e movimentação de pessoas, identifique uma eventual situação de emergência, a ideia é desenvolver um modelo preditivo que possa detectar tal comportamento. Nesse contexto, os modelos de classificação (como regressão múltipla, LASSO, Floresta Aleatória etc.) serão observados, verificando aquele com maior poder preditivo dentro do conjunto de dados utilizado.

O material compilado até aqui aponta caminhos para técnicas de detecção de anomalias, com base em *machine learning* (algoritmos como k-Nearest Neighbors e árvores de decisão, por exemplo) e *deep learning* (com o uso de redes neurais convolucionais - CNNs). Importante observar que mais de um artigo pesquisado tratou da questão das CNNs, que se mostram particularmente interessantes para detecção de anomalias devido à sua capacidade de lidar com dados não estruturados e aprender representações complexas.

O uso de imagens geolocalizadas para tomada de decisões durante as emergências também se mostra um caminho a ser considerado. Porém, talvez em um primeiro momento seja mais adequado começar com dados, gerados a partir de simulações,

que visem representar a movimentação das pessoas em determinadas situações. Dessa forma não haveria coleta de imagens, mas sim uma análise de dados fictícios gerados para simular situações reais.

E como o programa será desenvolvido em Python, será avaliada a biblioteca para detecção de anomalias em abordagens de aprendizado não supervisionado, desenhada para ser um ponto de referência na integração de técnicas de detecção de anomalias de última geração baseadas em grafos dinâmicos.

Considerando que o emprego da defesa NBQR por meio de Grupamentos Operativos de Fuzileiros Navais empresta flexibilidade para se adaptar exatamente à missão recebida, capacidade de planejamento e de logística própria, capacidade expedicionária e prontidão operativa, o desenvolvimento de uma inteligência artificial que possa identificar situações de emergência, em especial ataques químicos, seria de grande valia. O CFN está inserido na Estratégia Nacional de Defesa e sua capacitação em sistemas como o que pretende ser desenvolvido endereça aspectos estratégicos como a segurança da população, a proteção do patrimônio do país e a capacidade de resposta eficaz em situações anômalas.

### 3. METODOLOGIA

Com base nas possibilidades verificadas, verificou-se que a análise de dados com variáveis dependentes binárias, ou seja, aquelas que assumem apenas dois valores (0 ou 1), pode ser abordada por diferentes modelos de regressão. Para este experimento foram utilizados três modelos econométricos para estimar probabilidades e entender o comportamento de movimentação de pessoas em situações de emergência: o Modelo de Probabilidade Linear (MPL), o Modelo Probit e o Modelo Logit. Cada um desses métodos foi escolhido para abordar diferentes aspectos do fenômeno em questão e proporcionar uma análise abrangente. A seguir, descreve-se os detalhes da metodologia adotada, as características de cada modelo e a sua aplicabilidade no contexto do experimento.

#### 3.1 Modelo de Probabilidade Linear (MPL)

O MPL é uma técnica simples de regressão linear aplicada a variáveis dependentes binárias, estimando diretamente a probabilidade de um evento ocorrer. A relação entre a variável dependente e as variáveis independentes é expressa como uma combinação linear, ou seja:

$$P(Y = 1|X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

Onde  $Y$  é a variável dependente binária,  $X_1, X_2, \dots, X_k$  são as variáveis independentes, e  $\beta_0, \beta_1, \dots, \beta_k$  são os coeficientes estimados. Apesar de sua simplicidade e rapidez, o MPL pode gerar probabilidades preditas fora do intervalo  $[0, 1]$ , o que pode ser incoerente para modelagens probabilísticas.

Como vantagens podem ser citadas:

- 1) Simplicidade: Fácil de implementar e interpretar, especialmente em análises exploratórias;
- 2) Intuição: A relação linear entre as variáveis facilita a compreensão inicial dos dados; e

- 3) Rapidez: Pode ser computacionalmente mais eficiente do que modelos mais complexos.

Já entre as suas limitações:

- 1) Valores de probabilidade fora do intervalo [0, 1]: O MPL pode gerar probabilidades preditas menores que 0 ou maiores que 1, o que é incoerente com a interpretação de probabilidades;
- 2) Homoscedasticidade irrealista: Supõe que a variância do erro é constante ( $\sigma^2$ ), o que não é adequado para variáveis dependentes binárias;
- 3) Não-linearidade nas probabilidades reais: Relações complexas entre variáveis podem não ser bem representadas por um modelo linear; e
- 4) Substituição por modelos logit/probit: Modelos como a regressão logística (logit) ou probit são preferíveis em muitos casos, pois garantem que as probabilidades preditas estão dentro do intervalo válido e oferecem uma melhor modelagem da relação entre as variáveis.

Em relação à sua aplicação no experimento, o MPL foi utilizado para uma análise preliminar das variáveis que afetam a movimentação das pessoas. Ele ajudou a identificar padrões gerais e a gerar insights iniciais, mas suas limitações nas previsões de probabilidades tornaram necessárias outras abordagens.

### 3.2 Modelo Probit

O Modelo Probit é baseado na distribuição normal acumulada ( $\Phi$ ) e resolve algumas limitações do MPL ao garantir que as probabilidades previstas se mantenham dentro do intervalo [0, 1]. A fórmula do modelo é dada por:

$$P(Y = 1|X) = \Phi(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)$$

Onde  $\Phi$  é a função de distribuição acumulada normal e os coeficientes  $\beta_0, \dots, \beta_k$  são estimados via Máxima Verossimilhança (MLE).

Como vantagens podem- ser citadas:

- 1) As probabilidades preditas estão restritas ao intervalo [0, 1];
- 2) Modela bem relações não lineares entre a variável dependente e as variáveis independentes; ew
- 3) Adequado quando se assume que os erros seguem uma distribuição normal.

Entre suas limitações estão:

- 1) Maior complexidade computacional, especialmente na estimação via MLE; e
- 2) Interpretação dos coeficientes é mais difícil; os efeitos marginais devem ser calculados.

Em relação à sua aplicação no experimento, o Probit foi utilizado para modelar as probabilidades em contextos onde os dados seguem uma distribuição normal, como os deslocamentos regulares das pessoas. Este modelo proporcionou estimativas mais confiáveis em comparação ao MPL, especialmente em contextos menos extremos.

### 3.3 Modelo Logit

O Modelo Logit, assim como o Probit, é amplamente utilizado para modelar variáveis dependentes binárias. Ele usa a distribuição logística acumulada para garantir que as probabilidades preditas permaneçam dentro do intervalo [0, 1]. A fórmula do Logit é expressa como:

$$P(Y = 1|X) = \frac{1}{1+e^{-(\beta_0+\beta_1X_1+\dots+\beta_kX_k)}}$$

Onde  $e$  é a base do logaritmo natural e  $\beta_0, \dots, \beta_k$  são os coeficientes estimados.

Como vantagens podem- ser citadas:

- 1) Também garante que as probabilidades estejam dentro do intervalo válido [0, 1];

- 2) Modela bem relações não lineares; e
- 3) Interpretação direta em termos de *log-odds*, facilitando a compreensão dos efeitos das variáveis independentes.

Entre suas limitações estão:

- 1) Estimação mais complexa que o MPL, devido à necessidade de MLE; e
- 2) A assunção de distribuição logística pode ser inadequada em contextos onde outra distribuição (como a normal no Probit) seja mais apropriada.

Em relação à sua aplicação no experimento, o Logit foi utilizado especialmente em situações extremas ou de emergência, como alarmes ou bloqueios, onde as decisões binárias são fortemente influenciadas por fatores de alta variabilidade. Esse modelo é ideal para analisar esses eventos e prever as probabilidades de ocorrência de emergências, dado o impacto mais pronunciado das variáveis explicativas.

### **3.4 Integração dos Modelos**

Cada um dos modelos foi aplicado em diferentes estágios do experimento, aproveitando as suas características específicas:

O MPL foi utilizado para análises iniciais e rápidas, fornecendo uma visão geral da relação entre as variáveis. Ele foi útil para a exploração de padrões gerais e identificação de possíveis fatores de influência.

O Probit foi utilizado para modelar eventos relacionados a deslocamentos regulares, onde uma suposição de normalidade dos erros é mais apropriada. Ele garantiu que as probabilidades preditas estivessem dentro do intervalo  $[0, 1]$  e ofereceu uma modelagem mais precisa para contextos não lineares.

O Logit foi utilizado principalmente para modelar situações extremas, como emergências, onde as decisões binárias são mais influenciadas por eventos com maior variabilidade. Sua capacidade de modelar a relação entre variáveis de maneira

flexível foi fundamental para identificar padrões atípicos e acionar alertas de forma eficaz.

### **3.5 Ferramentas e Implementação**

Todos os modelos foram implementados utilizando o Python, que oferece suporte para estimar modelos como MPL (Modelo Linear de Probabilidade), Probit e Logit através de bibliotecas estatísticas robustas. Esses modelos são amplamente utilizados em econometria e aprendizado de máquina para prever a probabilidade de um evento.

Nesse contexto, o uso combinado de MPL, Probit e Logit permitiu uma análise abrangente e robusta do comportamento de movimentação das pessoas em diferentes contextos. O MPL foi útil para fornecer insights preliminares e análises rápidas, enquanto o Probit e o Logit garantiram maior precisão na modelagem de probabilidades e nas previsões de eventos, especialmente em contextos de deslocamento regular e emergencial. A escolha entre os modelos foi orientada pelas características dos dados e pelos objetivos do experimento, com a utilização do Python para garantir uma análise computacional eficiente e reprodutível.

## **4. DESENVOLVIMENTO DO TRABALHO**

Como mencionado no tópico 1.2 deste relatório, um modelo deve lidar com uma grande variação de densidades e movimentos presentes em uma multidão real para realizar este ambiente. É nesse contexto que a inteligência artificial a ser criada para esta finalidade tem por objetivo, a partir de dados de localização e movimentação de pessoas, responder informando se estar ocorrendo uma emergência ou não.

Para isto, foi definido neste trabalho um espaço fictício dentro de uma área militar que simula um local para o "rancho", ou seja, a refeição dos fuzileiros navais. Trata-se de um espaço estruturado com foco na funcionalidade, eficiência e, ao mesmo tempo, na disciplina. Ele será descrito a seguir.

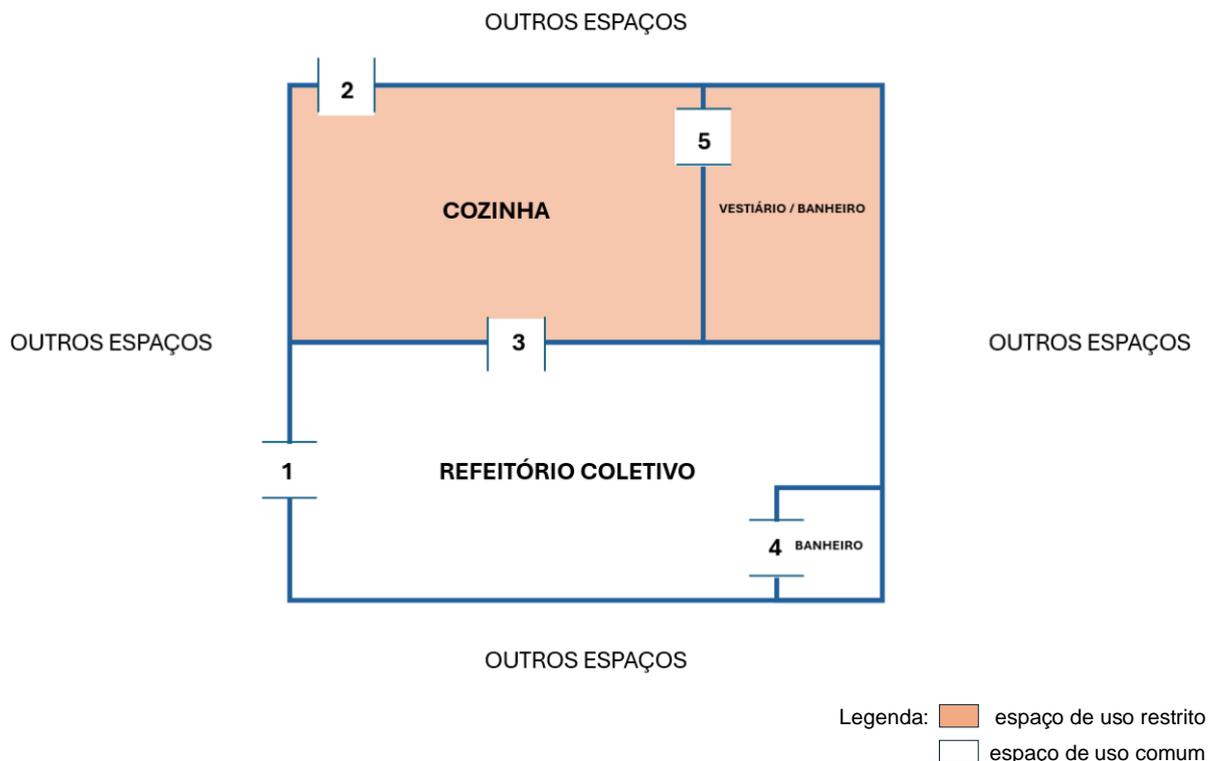
### **4.1 Descrição do espaço**

Do ponto de vista de leiaute, foi concebido um espaço de estrutura simples para observação do seu comportamento, com resposta, pela solução de inteligência artificial, sobre haver ou não uma emergência ocorrendo neste local em um dado momento. A área destinada ao rancho trata-se de uma grande sala ou um refeitório coletivo, projetado para comportar vários militares de uma só vez, mas com espaços bem definidos e organizados.

No espaço do refeitório há um banheiro de uso comum para os militares. Em anexo há uma cozinha para preparo das refeições que possui um vestiário/banheiro de uso exclusivo dos responsáveis pela cozinha. Importante registrar que este ambiente serve apenas almoço. O que não está inserido neste ambiente foi considerado neste modelo como "outros espaços".

A Figura 4.1.1 exemplifica o ambiente hipotético criado para este trabalho.

**Figura 4.1.1 - Espaço para a realização dos ranchos pelos militares**



Como é possível verificar na figura, o refeitório possui uma entrada/saída individual (acesso 1) e dois acessos à cozinha (acesso 2, vindo de “outros espaços”, e acesso 3, vindo do refeitório, ambos restritos à circulação dos funcionários dedicados a estes serviços). Podemos imaginar o refeitório com mesas longas e bancadas fixas, onde os fuzileiros navais fazem suas refeições. O refeitório coletivo possui um leiaute otimizado para que todos possam se servir de forma ágil e eficiente. As mesas são compostas por superfícies metálicas ou de plástico de fácil limpeza, e sobre elas geralmente há uma disposição padronizada para os utensílios, como pratos, talheres, copos e tigelas. Os utensílios são de metal e com o mesmo design simples e resistente utilizado em outros aspectos da vida militar. Em cada refeição o rancho segue um padrão, com porções distribuídas pelos funcionários da cozinha de maneira igualitária, de acordo com a quantidade de calorias e nutrientes necessários para as exigências físicas dos fuzileiros navais.

Importante observar que o rancho segue um horário fixo e rigoroso. Em muitas bases, os horários das refeições são anunciados de forma centralizada e há uma ordem de chamada para que os militares se dirijam ao refeitório de maneira disciplinada. Para

evitar desorganização, o serviço de alimentação é altamente eficiente, com a distribuição das porções controlada. Nenhum fuzileiro naval pode sair do refeitório sem completar sua refeição e não é permitido fazer "segundas porções" de forma arbitrária.

O ambiente é funcional, sem muitos elementos decorativos, refletindo a filosofia militar de eficiência e pragmatismo. Há, no entanto, algumas placas ou bandeiras nas paredes, com símbolos militares, ou outros emblemas relacionados à instituição. Há quadro de avisos com informações sobre os horários das atividades diárias ou instruções importantes. Isto significa que a circulação é ágil e sem impedimentos de qualquer natureza neste ambiente.

Dentro do refeitório há um banheiro de uso comum, simples, funcional e de fácil acesso (acesso 4), mantido em condições adequadas de uso, com um único acesso ao refeitório. O clima dentro do rancho é de respeito, com os militares se locomovendo de forma tranquila e organizada. Enquanto as refeições são tomadas, os fuzileiros navais podem conversar, mas sempre dentro dos limites do respeito e da disciplina, com a maior parte da atenção voltada à comida e ao momento da refeição em si.

Após o término da refeição, todos os militares têm a obrigação de deixar a mesa limpa. As regras de etiqueta são rigorosas: o fuzileiro naval deve organizar seus utensílios e deixá-los na mesa para que o serviço de limpeza possa ser feito de forma eficiente.

Anexa ao refeitório, há uma cozinha, ambiente restrito e altamente organizado, dedicado exclusivamente à preparação das refeições para os fuzileiros navais. A cozinha é composta por uma série de equipamentos industriais e profissionais, como fogões grandes, fornos, fritadeiras e recipientes de armazenamento. Tudo é mantido de acordo com as normas de higiene e segurança alimentar. Importante registrar que a cozinha possui um acesso externo individual (acesso 2), por onde os seus funcionários, tanto aqueles que cuidam das refeições quanto os que tratam da limpeza do local, a acessam, sem necessidade de trânsito pela entrada/saída externa do refeitório, destinada basicamente àqueles militares que se dirigem ao ambiente para fazer suas refeições.

A cozinha é separada fisicamente do refeitório por uma única porta (acesso 3) por onde são servidas as refeições aos militares no refeitório. Sendo de acesso restrito aos funcionários responsáveis pela alimentação, existe um fluxo de trabalho altamente organizado, em que cada etapa da preparação é executada por diferentes equipes para garantir que a comida seja entregue no momento certo e com a qualidade necessária.

Em um espaço anexo à cozinha há um vestiário/banheiro exclusivo para os seus funcionários (acesso 5). Essa área é acessada por uma única entrada/saída e é restrita ao pessoal que trabalha na preparação das refeições. Nela os cozinheiros e assistentes se trocam, higienizam-se e descansam brevemente entre os turnos de trabalho.

Trata-se de um ambiente simples, porém funcional, com armários individuais para que cada funcionário possa guardar suas roupas pessoais e equipamentos de trabalho. O ambiente é projetado para garantir a segurança e higiene dos que trabalham na cozinha. Há chuveiros e sanitários limpos e organizados, com um controle rigoroso sobre a limpeza e manutenção do local.

Mais uma vez observa-se que a entrada tanto para a cozinha quanto para o vestiário/banheiro dos funcionários é restrita. Apenas aqueles com funções específicas dentro da equipe de cozinha têm permissão para acessar essas áreas. Isso é feito para preservar a higiene, segurança e eficiência do trabalho, além de evitar distrações ou desordem durante a preparação das refeições.

Para efeito de designação no código python desenvolvido, os espaços são assim identificados (não há acentos e caracteres especiais, pois reproduzem exatamente como estão identificados no código python):

#local: complexo de alimentacao da marinha (so serve almoco)

#espaco 1111: outros espacos

#espaco 2222: cozinha

#espaco 3333: refeitorio

#espaco 4444: banheiro

#espaco 5555: vestiario/banheiro

## 4.2 Descrição dos atores envolvidos na observação do modelo

No espaço descrito no tópico 2.1 estão envolvidos atores cujos comportamentos serão observados pelo modelo de inteligência artificial proposto. Para cada tipo de indivíduo serão considerados 3 (três) tipos de comportamento (para efeito de designação no código python desenvolvido, os espaços são assim identificados - não há acentos e caracteres especiais, pois reproduzem exatamente como estão identificados no código python):

#comportamento tradicional ou rotineiro

#comportamento atipico (doenca ou ferias)

#comportamento de emergencia (situacao de emergencia)

Os tipos de indivíduos são os comensais (os fuzileiros navais que entram no espaço de estudo para almoçar), os cozinheiros (fuzileiros navais que trabalham no local e são divididos em dois tipos distintos) e o pessoal de limpeza (fuzileiros que também trabalham no local e se dividem em dois tipos diferentes). Eles serão descritos a seguir.

Como informação final o período do modelo considera dias úteis de 2020 a 2023 (1.014 dias úteis) e número de dias de emergência total de 40 (média de 10 por ano).

### 4.2.1 Comensais (#tipo 1000)

São chamados de “comensais” aqueles fuzileiros navais que se dirigem ao refeitório para fazerem suas refeições. Eles são do tipo 1000 e correspondem a um total de 300 (trezentos) indivíduos, que chegam ao local entre 11h30 e 12h30 para o almoço. Neste intervalo eles acessam o refeitório a partir de outros espaços da base militar pelo acesso 1.

No modelo estatístico é feito um sorteio por distribuição uniforme. O horário de entrada de cada comensal dentro da janela de 11h30 a 12h30 é, portanto, aleatoriamente

determinado, com cada horário dentro deste intervalo tendo a mesma probabilidade de ser sorteado.

Por exemplo, no dia inicial da observação, um comensal pode ter o horário de entrada de 11h34 e, no dia seguinte, o horário sorteado pode ser diferente, 11h45. O modelo proposto realiza este sorteio para cada dia, garantindo que a distribuição de horários seja aleatória e sem qualquer padrão previsível.

Na sequência, para o mesmo comensal, o modelo faz um outro sorteio, também numa variável uniforme, considerando um intervalo de 20 a 45 minutos, para saber quanto tempo ele vai permanecer no refeitório. Isto significa que o modelo considera que ele permanece no refeitório almoçando entre 20 (mínimo) e 45 minutos (máximo). Com isso é possível saber que horas aquele comensal irá sair do refeitório. Como um exemplo, se o sorteio indica 30 minutos, isto significa que ele fez sua refeição neste tempo e saiu do refeitório.

O modelo considera ainda que o comensal, obrigatoriamente, irá ao banheiro de acesso comum depois de fazer a sua refeição. Neste caso é realizado um novo sorteio de uma variável uniforme aleatória, em um intervalo entre 5 e 15 minutos, mínimo e máximo estabelecidos para o uso do banheiro. Com isso é possível saber a hora que o comensal entrou no banheiro e o momento em que ele saiu. Ao sair do banheiro ele volta obrigatoriamente para o refeitório, uma vez que a saída é pelo mesmo acesso 1 de entrada no refeitório.

Neste momento é feito um novo sorteio de uma variável aleatória uniforme entre 2 e 10 minutos. Este sorteio avalia a probabilidade do comensal sair do banheiro direto para fora do refeitório (2 minutos) ou encontrar algum conhecido e conversar rapidamente (até 10 minutos).

Com isso o modelo obtém exatamente o horário de entrada e saída do comensal no refeitório. Somando-se o tempo máximo de todos os sorteios que são feitos, quais sejam, o tempo para almoçar (45 minutos), o tempo de uso do banheiro (15 minutos) e o tempo do banheiro até a saída (10 minutos), chega-se ao limite máximo de 70 minutos que o comensal pode ficar neste espaço. Sendo assim, se um comensal

entrou 12h30 no refeitório sua saída poderá ocorrer, no máximo, às 13h40. Nesse contexto, estabeleceu-se para o modelo o horário de funcionamento do refeitório entre 11h30 e 13h45.

Esta é a rotina do comensal, seu comportamento padrão. No caso de uma emergência há algumas possibilidades que o modelo considera para este indivíduo. Se ocorrer, por exemplo, uma emergência antes do horário de abertura da refeição, o modelo considera que o comensal não sairá de outros espaços para entrar no refeitório. Se a emergência ocorrer no intervalo entre 11h30 e 12h30, somente alguns dos 300 comensais observados não entrarão, portanto, nem todos os sorteios serão realizados.

Neste último caso, ocorrendo uma emergência o comensal tem um tempo de resposta que varia de 0 a 5 minutos. Se ele estiver no refeitório é feito um sorteio neste intervalo para que se dirija a outros espaços pelo acesso 1. Se o comensal estiver no banheiro é feito um sorteio neste intervalo para que ele se dirija para o refeitório e de lá para outros espaços pelo acesso 1, lembrando que o modelo, em caso de emergência, busca sempre o acesso de saída mais rápido deste espaço. Finalmente, se no momento da emergência o comensal que almoçou naquele dia já tiver saído do refeitório para outros espaços, o modelo desconsidera este fato, pois a rotina dele foi normal durante o tempo em que esteve neste ambiente.

Como já mencionado, para efeito de designação no código python desenvolvido, os 300 (trezentos) comensais foram identificados como “#tipo 1000”.

#### **4.2.2 Cozinheiros**

São chamados de “cozinheiros” aqueles fuzileiros navais que se dirigem ao espaço de observação do modelo para preparar e servir as refeições. Eles correspondem, portanto, a dois tipos: aqueles que permanecem na cozinha fazendo as refeições (#tipo 2000) e aqueles que transitam entre a cozinha e o refeitório, repondo a comida e servindo a fila de comensais (#tipo 3000). Eles compõem um total de 20 (vinte) indivíduos, sendo 15 (quinze) do tipo 2000 e 5 (cinco) do tipo 3000.

#### **4.2.2.1 Cozinheiros (#tipo 2000)**

Os cozinheiros do tipo 2000, num total de 15 (quinze) indivíduos, chegam ao local vindo de “outros espaços” entre 6h30 e 7h30 da manhã. Neste intervalo eles acessam a cozinha pelo acesso 2. Eles chegam bem cedo porque há um volume muito grande de refeições a serem preparadas e isso demanda um esforço grande desta equipe para que às 11h30 tudo esteja pronto no refeitório.

No modelo eles estão submetidos à uma distribuição aleatória uniforme que sorteia qualquer número nesse intervalo com a mesma probabilidade entre 6h30 e 7h30 da manhã de entrada neste espaço. Uma vez determinado o horário de sua entrada, sua permanência inicial na cozinha é rápida, pois ele deve se dirigir ao vestiário para se trocar (pelo acesso 5). Assim, é feito um novo sorteio, de maneira uniforme, que determinará esta permanência entre 2 a 5 minutos até que ele se dirija ao vestiário.

Dessa forma, o modelo já terá estimado o horário que o cozinheiro saiu de outros espaços e entrou na cozinha, o tempo que permaneceu ali e o horário que entra no vestiário. No vestiário o modelo vai considerar um sorteio entre 15 e 30 minutos de permanência no vestiário, considerando a possibilidade do indivíduo não apenas se trocar, mas eventualmente tomar um banho, usar o banheiro etc. O sorteio, mais uma vez, é feito de maneira uniforme e aleatória, com a probabilidade de ocorrência de qualquer valor.

Quando este cozinheiro sai do vestiário e volta para a cozinha, ele começa efetivamente a trabalhar. O modelo estima que ele irá preparar as refeições entre 150 e 180 minutos. Neste intervalo há uma série de atividades sendo realizadas como o preparo das refeições, recebimento de mantimentos, armazenamento do material recebido, conferência de alimentos vencidos na dispensa e/ou freezers e geladeiras, enfim, uma gama de atividades inerentes a este perfil. Mais uma vez o sorteio é realizado pelo modelo dentro de uma maneira uniforme e aleatória, sorteando qualquer número com a mesma probabilidade entre 150 e 180 minutos.

Estas atividades ocorrem antes dos comensais chegarem, pois, no pior caso do sorteio, um cozinheiro teria um sorteio em que suas atividades iriam até às 10h30

(entrada às 7h30 mais 180 minutos de trabalho). Inclusive o modelo assume que estes indivíduos possam almoçar neste período, antes da chegada dos comensais, pois eles chegam muito cedo e, eventualmente, podem estar com fome mais cedo que 11h30, horário inicial da entrada dos comensais. Mais uma vez o modelo considera uma ida ao vestiário após a preparação da refeição (entre 15 e 30 minutos), também com um sorteio uniforme e aleatório desta permanência, voltando para a cozinha e permanecendo ali até o final do seu horário de trabalho, colaborando na logística de reposição da comida e organizando tudo no seu ambiente de trabalho.

A conta realizada pelo modelo, em resumo, é a seguinte: a partir do horário que ele entrou na cozinha (sorteado), este indivíduo trabalhará 7h45. Os 15 minutos de sobra são para a sua ida ao vestiário, no final de sua jornada, para se trocar e voltar para a casa. 15 minutos é o tempo mínimo nesta ida final ao vestiário, mas o sorteio que é feito neste momento também considera a possibilidade de ele utilizar até 30 minutos para sair do vestiário e voltar para a cozinha (passagem obrigatória) antes de sair definitivamente para outros espaços. Nesta passagem final pela cozinha é realizado um novo sorteio (de maneira uniforme e aleatório, entre 2 e 5 minutos), que é o tempo em que ele se despede dos colegas etc.

Então essa é a rotina do cozinheiro. No caso de uma emergência o modelo trata da mesma forma como já relatado para o comensal. Há um tempo de resposta que varia de 0 a 5 minutos. Cabe registrar que o sorteio nunca irá considerar um tempo de resposta igual a 0, pois há obrigatoriamente um atraso nesta resposta. Assim, se ele estiver na cozinha é feito um sorteio neste intervalo para que se dirija a outros espaços pelo acesso 2. Se o cozinheiro estiver no vestiário é feito um sorteio neste intervalo para que se dirija à cozinha e de lá saia para outros espaços pelo acesso 2, lembrando que o modelo, em caso de emergência, busca sempre o acesso de saída mais rápido de cada espaço. Finalmente, se no momento da emergência o cozinheiro não tiver entrado vindo de outros espaços ou já tiver saído da cozinha para outros espaços, o modelo desconsidera este fato, pois a rotina dele foi normal durante o tempo em que esteve neste ambiente.

O horário de saída deste tipo de cozinheiro é, no máximo, 15h45.

Como já mencionado, para efeito de designação no código python desenvolvido, os 15 (quinze) cozinheiros foram identificados como “#tipo 2000”.

#### **4.2.2.2 Cozinheiros (#tipo 3000)**

Os cozinheiros do tipo 3000, num total de 5 (cinco) indivíduos, não servem diretamente os comensais, mas recolhem as bandejas para lavagem, repõe a comida nas gôndolas, fazem a operação do serviço de almoço. Podem eventualmente recolher bandejas esquecidas nas mesas, jogar resto de comida fora, limpar a mesa (ainda que não sejam o pessoal da limpeza).

A sua rotina diária é bem semelhante ao cozinheiro tipo 2000. Eles chegam ao local vindo de “outros espaços” entre 6h30 e 7h30 da manhã. Neste intervalo eles acessam a cozinha pelo acesso 2. O modelo assume que, neste período, que os comensais ainda não chegaram, estes cozinheiros dão apoio aos cozinheiros do tipo 2000, ajudando a preparar as refeições, organizar os equipamentos etc.

Então o comportamento padrão ocorre da mesma forma que o outro cozinheiro. No modelo eles estão submetidos à uma distribuição aleatória uniforme que sorteia qualquer número nesse intervalo com a mesma probabilidade entre 6h30 e 7h30 da manhã de entrada neste espaço. Uma vez determinado o horário de sua entrada, sua permanência inicial na cozinha é rápida, pois ele deve se dirigir ao vestiário para se trocar (pelo acesso 5). Assim, é feito um novo sorteio, de maneira uniforme, que determinará esta permanência entre 2 a 5 minutos até que ele se dirija ao vestiário.

Dessa forma, o modelo já terá estimado o horário que o indivíduo saiu de outros espaços e entrou na cozinha, o tempo que permaneceu ali e o horário que entra no vestiário. No vestiário o modelo vai considerar um sorteio entre 15 e 30 minutos de permanência no vestiário, considerando a possibilidade do indivíduo não apenas se trocar, mas eventualmente tomar um banho, usar o banheiro etc. O sorteio, mais uma vez, é feito de maneira uniforme e aleatória, com a probabilidade de ocorrência de qualquer valor.

Quando este cozinheiro sai do vestiário e volta para a cozinha, ele começa efetivamente a trabalhar. O modelo estima que ele irá ajudar os cozinheiros do tipo 2000 no preparo das refeições entre 150 e 180 minutos. Para este intervalo o sorteio é realizado pelo modelo dentro de uma maneira uniforme e aleatória, sorteando qualquer número com a mesma probabilidade entre 150 e 180 minutos. Da mesma este indivíduo também almoça mais cedo, se dirige ao vestiário, permanece ali entre 15 e 30 minutos (sorteio de maneira uniforme e aleatória), passa rapidamente pela cozinha (novo sorteio do modelo entre de 2 e 5 minutos), para finalmente dirigir-se ao refeitório pelo acesso 3 e aí ele vai pro refeitório, pronto para atender os comensais.

Daí em diante ele passa o restante do dia basicamente no refeitório. Da mesma forma que realizado para o cozinheiro tipo 2000, a partir do momento em que ele entra pela manhã na cozinha (sorteado), este indivíduo trabalhará 7h45. Os 15 minutos de sobra são para a sua ida ao vestiário, no final de sua jornada, para se trocar e voltar para a casa. 15 minutos é o tempo mínimo nesta ida final ao vestiário, mas o sorteio que é feito neste momento também considera a possibilidade de ele utilizar até 30 minutos para sair do vestiário e voltar para a cozinha (passagem obrigatória) antes de sair definitivamente para outros espaços. Nesta passagem final pela cozinha é realizado um novo sorteio (de maneira uniforme e aleatório, entre 2 e 5 minutos), que é o tempo em que ele se despede dos colegas etc.

Então essa é a rotina do cozinheiro tipo 3000. No caso de uma emergência o modelo trata da mesma forma como já relatado para os demais indivíduos. Há um tempo de resposta que varia de 0 a 5 minutos (Obs: como já mencionado, o sorteio nunca irá considerar um tempo de resposta igual a 0, pois há obrigatoriamente um atraso nesta resposta). Assim, se ele estiver na cozinha é feito um sorteio neste intervalo para que se dirija a outros espaços pelo acesso 2. Se o cozinheiro estiver no vestiário é feito um sorteio neste intervalo para que se dirija à cozinha pelo acesso 5 e de lá saia para outros espaços pelo acesso 2, lembrando que o modelo, em caso de emergência, busca sempre o acesso de saída mais rápido de cada espaço. Adicionalmente, em uma emergência o cozinheiro do tipo 3000 poderá estar no refeitório e, considerando este fato, ele deve procurar a saída mais rápida, direto pelo acesso 1. Ele é o único tipo de cozinheiro que tem a opção de sair por dois acessos distintos, dependendo de sua localização no momento do evento emergencial.

Finalmente, se no momento da emergência o cozinheiro não tiver entrado vindo de outros espaços ou já tiver saído da cozinha ou do refeitório para outros espaços, o modelo desconsidera este fato, pois a rotina dele foi normal durante o tempo em que esteve neste ambiente.

O horário de saída deste tipo de cozinheiro é, no máximo, 15h45.

Como já mencionado, para efeito de designação no código python desenvolvido, os 5 (cinco) cozinheiros foram identificados como “#tipo 3000”.

### **4.2.3 Pessoal da limpeza**

São chamados de “pessoal da limpeza” aqueles fuzileiros navais que se dirigem ao espaço de observação do modelo para fazer a limpeza do local das refeições. Eles correspondem, portanto, a dois tipos: aqueles chegam mais tarde (#tipo 4000) e aqueles que chegam mais cedo (#tipo 5000). Eles compõem um total de 7 (sete) indivíduos, sendo 5 (cinco) do tipo 4000 e 2 (dois) do tipo 5000.

#### **4.2.3.1 Pessoal da Limpeza (#tipo 4000)**

O pessoal da limpeza do tipo 4000, num total de 5 (cinco) indivíduos, são responsáveis por limpar o ambiente todo (cozinha, refeitórios e banheiro) após a saída de todos os comensais e cozinheiros.

A sua rotina diária é chegar ao local vindo de “outros espaços” entre 16h00 e 16h30 da tarde. Neste intervalo eles acessam a cozinha pelo acesso 2, submetidos à uma distribuição aleatória uniforme que sorteia qualquer número nesse intervalo com a mesma probabilidade entre 16h00 e 16h30 para registro de entrada neste espaço. Uma vez determinado o horário de sua entrada, sua permanência na cozinha é de 30 a 60 minutos, sorteados de maneira uniforme e aleatória. Aqui este indivíduo fará a lavagem do chão, recolher lixo, entre outras atividades de limpeza do local. Se o tempo de permanência é menor, assume-se que a situação para limpeza era melhor. O pior caso leva a uma permanência maior.

Na sequência passam de 30 a 60 minutos no vestiário (entrando pelo acesso 5), sorteado de maneira uniforme e aleatória. Limpam o local, repõe sabonetes líquidos, toalhas de papel etc. Voltam para a cozinha e, como os outros indivíduos, tem um sorteio uniforme e aleatório de sua permanência ali entre 2 e 5 minutos.

Na sequência passam para o refeitório pelo acesso 3, passando de 30 a 60 minutos naquele local, sorteados com base em uma variável aleatória uniforme. Li limpam as mesas, o chão, retiram o lixo das lixeiras, limpas as paredes (se couber), deixando tudo apto a utilização no dia seguinte.

Dali passam para o banheiro pelo acesso 4, podendo levar entre 20 e 45 minutos limpando o local (sorteados de maneira uniforme e aleatória). Voltam ao refeitório, passam de 2 a 5 minutos ali (sorteados de maneira uniforme e aleatória) e vão embora, saindo diretamente pelo acesso 1.

Essa é a rotina do indivíduo do tipo 4000. No caso de uma emergência o modelo trata da mesma forma como já relatado para os demais indivíduos. Há um tempo de resposta que varia de 0 a 5 minutos. Assim, se ele estiver na cozinha é feito um sorteio neste intervalo para que se dirija a outros espaços pelo acesso 2. Se ele estiver no vestiário é feito um sorteio neste intervalo para que se dirija à cozinha pelo acesso 5 e de lá saia para outros espaços pelo acesso 2, lembrando que o modelo, em caso de emergência, busca sempre o acesso de saída mais rápido de cada espaço. Adicionalmente, em uma emergência, se ele estiver no refeitório deverá procurar a saída mais rápida, direto pelo acesso 1. Caso esteja no banheiro do refeitório, sai pelo acesso 4 ao refeitório e dali para outros espaços via acesso 1.

Como nas outras situações, se no momento da emergência este indivíduo não tiver entrado vindo de outros espaços ou já tiver saído deste ambiente para outros espaços, o modelo desconsidera este fato, pois a rotina dele foi normal durante o tempo em que esteve neste ambiente.

Considerando a jornada de 8 horas, o horário máximo de permanência no ambiente de observação é até 00h30.

Como já mencionado, para efeito de designação no código python desenvolvido, as 5 (cinco) pessoas da limpeza foram identificadas como “#tipo 4000”.

#### **4.2.3.2 Pessoal da Limpeza (#tipo 5000)**

Por fim, o último tipo do pessoal da limpeza, 5000, num total de 2 (dois) indivíduos, são responsáveis por manter a cozinha limpa enquanto os cozinheiros ainda estão trabalhando, retirando lixo, comida no chão etc. Além disso, mantém limpo o vestiário, o refeitório (limpando o chão de algo que entornou, por exemplo) e o banheiro do refeitório, para que os comensais possam utilizá-lo sem problemas.

A sua rotina diária é chegar ao local vindo de “outros espaços” entre 11h00 e 11h30 da manhã, antes da chegada dos comensais. Neste intervalo eles acessam a cozinha pelo acesso 2, submetidos à uma distribuição aleatória uniforme que sorteia qualquer número nesse intervalo com a mesma probabilidade entre 11h00 e 11h30 para registro de entrada neste espaço. Uma vez determinado o horário de sua entrada, sua permanência na cozinha é de 30 a 45 minutos, sorteados de maneira uniforme e aleatória. Aqui, como já mencionado, este indivíduo fará a manutenção da limpeza do local para os cozinheiros, recolhendo lixo, entre outras atividades de limpeza do local.

Na sequência passam de 15 a 25 minutos no vestiário (entrando pelo acesso 5), sorteado de maneira uniforme e aleatória. Mantém o local limpo, volta para a cozinha e, como os outros indivíduos, tem um sorteio uniforme e aleatório de sua permanência ali entre 2 e 5 minutos.

Na sequência passam para o refeitório pelo acesso 3, passando de 30 a 45 minutos naquele local, sorteados com base em uma variável aleatória uniforme. Ali limpam as mesas, o chão, mantendo o local adequado enquanto os comensais estão fazendo suas refeições.

Dali passam para o banheiro pelo acesso 4, mas aqui ficam pouco tempo, entre 5 e 15 minutos limpando o local (sorteados de maneira uniforme e aleatória). Eles não ficam muito tempo ali porque o banheiro está em uso, os comensais estão no local.

Voltam ao refeitório, passam de 2 a 5 minutos ali (sorteados de maneira uniforme e aleatória) e vão embora, saindo diretamente pelo acesso 1.

Essa é a rotina do indivíduo do tipo 5000. No caso de uma emergência o modelo trata da mesma forma como já relatado para os demais indivíduos. Há um tempo de resposta que varia de 0 a 5 minutos. Assim, se ele estiver na cozinha é feito um sorteio neste intervalo para que se dirija a outros espaços pelo acesso 2. Se ele estiver no vestiário é feito um sorteio neste intervalo para que se dirija à cozinha pelo acesso 5 e de lá saia para outros espaços pelo acesso 2, lembrando que o modelo, em caso de emergência, busca sempre o acesso de saída mais rápido de cada espaço. Adicionalmente, em uma emergência, se ele estiver no refeitório deverá procurar a saída mais rápida, direto pelo acesso 1. Caso esteja no banheiro do refeitório, sai pelo acesso 4 ao refeitório e dali para outros espaços via acesso 1.

Como nas outras situações, se no momento da emergência este indivíduo não tiver entrado vindo de outros espaços ou já tiver saído deste ambiente para outros espaços, o modelo desconsidera este fato, pois a rotina dele foi normal durante o tempo em que esteve neste ambiente.

Considerando a jornada de 8 horas, o horário máximo de permanência no ambiente de observação é até 19h30.

Como já mencionado, para efeito de designação no código python desenvolvido, as 2 (duas) pessoas da limpeza foram identificadas como “#tipo 5000”.

## 5. DESCRIÇÃO DA ROTINA

Com base no espaço de observação concebido e nos atores participantes, descritos no capítulo 4, todos devidamente caracterizados, foi estruturada a rotina do modelo que vai simular este fenômeno e responder a uma eventual emergência com base no comportamento dos atores envolvidos.

### 5.1 Primeiras linhas da rotina

As primeiras linhas da rotina em Python estão apresentadas a seguir:

```
#Deteccao de emergencia #####  
#Explicacoes #####  
  
#local: complexo de alimentacao da marinha (so serve almoco)  
#espaco 1111: outros espacos  
#espaco 2222: cozinha  
#espaco 3333: refeitório  
#espaco 4444: banheiro  
#espaco 5555: vestiario/banheiro  
  
#pessoas: cinco tipos  
#tipo 1000: comensais (300 desse tipo)  
#tipo 2000: cozinheiros - ficam na cozinha / cozinhando (15 desse tipo)  
#tipo 3000: cozinheiros - ficam no refeitório / repondo comida / servindo a fila (5 desse tipo)  
#tipo 4000: pessoal da limpeza / fim do dia (5 desse tipo)  
#tipo 5000: pessoal da limpeza / durante o dia (2 desse tipo)  
  
#para cada tipo de pessoa serao considerados tres comportamentos  
#comportamento tradicional ou rotineiro  
#comportamento atipico (doenca ou ferias)  
#comportamento de emergencia (situacao de emergencia)  
  
#periodo do modelo: dias uteis de 2020 a 2023 (1014 dias uteis)  
#numero de dias de emergencia: 40 dias de emergencia (media de 10 por ano)
```

Aqui são definidos os ambientes de observação (outros espaços, cozinha, refeitório, banheiro, vestiario/banheiro), como já apresentado no item 2.2, os cinco tipos de

atores (comensais – tipo 1000; cozinheiros – Tipo 2000; cozinheiros – tipo 3000; pessoal da limpeza – tipo 4000; e pessoal da limpeza – tipo 5000).

Além disso, para cada tipo de ator (ou pessoa) observada serão considerados três comportamentos:

- 1) comportamento tradicional ou rotineiro;
- 2) comportamento atípico (doença ou férias); e
- 3) comportamento de emergência (situação de emergência).

O comportamento atípico, que considera doença ou férias, caracteriza uma eventual ausência mais prolongada de um dos atores nos dias observados, sem, contudo, caracterizar algum tipo de emergência.

Por fim, o período adotado no modelo considera dias úteis entre 2020 e 2023, totalizando 1.014 dias úteis. E o número de dias caracterizados como “emergência” serão 40 (quarenta) para o período, média de 10 (dez) por ano.

## **5.2 Instalação de pacotes**

Na sequência foi feito o processo de instalação de todos os pacotes necessários. Esses pacotes são essenciais para a modelagem e simulação dos dados e para o correto funcionamento da rotina do código. São eles: pandas; holidays; numpy; e random.

## **5.3 Criação da base de dados**

Em seguida foi feita a criação da base de dados. O primeiro passo foi construir um vetor que contém todos os dias úteis dos anos de 2020 a 2023, ou seja, quatro anos de dias úteis. Esse vetor foi gerado automaticamente para incluir apenas os dias úteis de cada ano (excluindo fins de semana e feriados).

Após isso, foram criadas as datas de emergência e as horas em que as emergências ocorrem. Para isso foi feito um sorteio aleatório a partir do vetor de dias úteis já gerado.

Escolheu-se, assim, 40 dias, entre os dias úteis, sorteada também a hora da emergência, que pode ocorrer entre 7h da manhã e 20h da noite. Para cada dia sorteado, também foi sorteada a duração da emergência, que pode ser de 1 a 12 horas. Esse tempo de duração foi adicionado ao horário da emergência. Por exemplo, se a emergência começa às 7h e a duração sorteada for 3 horas, então ela durará até as 10h.

Se a emergência ultrapassar o final do dia, por exemplo, começando às 18h, e durar 5 horas, ela será "travada" até meia-noite, para refletir a realidade de que emergências fora do horário de trabalho padrão são menos comuns nesta modelagem.

A escolha de um período entre 7h e 20h para as emergências é estratégica, pois o que está sendo modelado é um ambiente de trabalho, onde os indivíduos são mais propensos a estarem presentes durante esse intervalo de tempo.

#### **5.4 Criação dos indivíduos**

A criação dos indivíduos que irão compor a base de dados foi feita na sequência. Para simplificar, criou-se inicialmente 300 indivíduos do tipo **\*\*1000\*\*** (comensais), para os quais sua rotina e comportamento serão modelados.

Primeiro foi criado um *DataFrame* vazio para alocar esses indivíduos. Utilizou-se um *loop* (for) de 1 até 300 para criar 300 indivíduos e, para cada um deles, sorteou-se a sua rotina diária. Isso inclui, entre outras variáveis, a hora de entrada e saída de cada local que ele frequenta.

A rotina de cada indivíduo foi definida da seguinte maneira: a) O indivíduo começa em "outros espaços" (o local onde ele inicia o dia). A hora em que ele entra e sai desse espaço foi sorteada, assim como a velocidade com que ele se locomove entre os espaços; b) o indivíduo segue para o "refeitório", e, assim, sucessivamente, para os outros espaços que frequenta, como vestiário/banheiro etc.

Esse processo é repetido até obter-se todas as variáveis de tempo (hora de entrada, hora de saída) e de movimento (velocidade de locomoção) definidas para todos os

espaços que o indivíduo frequenta. A partir disso é criado o comportamento diário de cada indivíduo, com todas essas informações.

## **5.5 Alteração da rotina devido a emergências, doença e férias**

Com os comportamentos diários dos indivíduos prontos, começou-se, então, a ajustar essas rotinas conforme eventos extraordinários como emergências, dias de doença e férias, descritas a seguir.

**Emergências** - Quando ocorrer uma emergência em um dia específico, ajusta-se a rotina do indivíduo para refletir o impacto da emergência. Por exemplo, se o indivíduo estava no refeitório quando a emergência começou, sua rotina a partir daquele momento será alterada.

Em relação à emergência, três principais mudanças ocorrem:

- 1) A velocidade de locomoção do indivíduo aumenta, pois ele começa a correr (passando de 4 a 6 km/h para 9 a 12 km/h);
- 2) O trajeto do indivíduo é alterado, ele deixa de seguir sua rotina normal para seguir um novo trajeto devido à emergência; e
- 3) As horas de entrada e saída também mudam. O indivíduo pode precisar sair antes ou depois dos horários tradicionais devido à ocorrência de emergência.

**Doença** - Para cada indivíduo é sorteado dois dias de doença por ano. Nos dias de doença a rotina do indivíduo é completamente alterada, pois ele não aparece para trabalhar. Esses dias de doença podem coincidir com dias de emergência ou férias. Se isso acontecer, o comportamento de rotina do indivíduo é substituído pelo comportamento de doença. O sorteio de dias de doença é feito considerando que o indivíduo terá no máximo 8 dias de doença ao longo de 4 anos (2 dias por ano).

**Férias** - Cada indivíduo também tem direito a dois períodos de férias de 15 dias corridos, que podem ser sorteados para ocorrer no início ou meio do ano. Durante as férias o indivíduo não aparece para trabalhar, o que afeta diretamente sua rotina.

Como as férias são contadas como 15 dias corridos, o número de dias úteis (excluindo fins de semana) será menor, normalmente 11 dias úteis.

Para exemplificar, se a emergência acontece e o indivíduo estava no refeitório, tudo o que ele faz até chegar neste espaço é normal e o seu comportamento muda exatamente no refeitório. Isso acaba ocupando um espaço enorme da rotina, são muitos “if” para a hora da emergência: se o indivíduo estiver no espaço 1 ele vai ter um comportamento, se estiver no espaço 2 é outro e, assim, sucessivamente. São muitas linhas da rotina.

Cabe observar que, em dias de doença, o comportamento também é alterado. São sorteados dois dias de doença, onde ele não comparece ao trabalho. Cabe registrar que um dia de doença pode ser um dia de emergência. Se isto ocorrer, para o comportamento de emergência deste indivíduo que poderia estar no refeitório é feito um *override*. Se estiver em férias, ocorre a mesma coisa. São sorteados alguns dias para férias e também é feito um *override*, com o indivíduo ausente do espaço.

## **5.6 Modelagem de outros tipos de indivíduos**

Além dos “comensais” (tipo 1000), foram modelados, como já informado, outros tipos de indivíduos, como “cozinheiros” e “pessoal da limpeza”.

**Indivíduos do tipo 2000 (cozinheiros)** – Criou-se 15 indivíduos desse tipo e, para cada um deles, foi definida a sua rotina diária. A principal diferença é que os cozinheiros passam o dia inteiro na cozinha e sua rotina é adaptada para refletir esse comportamento. Eles têm uma rotina mais fixa, com menos variação de espaço se comparado aos comensais.

Assim, como no caso dos comensais, a rotina dos cozinheiros é ajustada em caso de emergência, doença ou férias.

**Indivíduos do tipo 3000 (cozinheiros que também transitam pelo refeitório)** - Esses indivíduos frequentam mais espaços do que os cozinheiros do tipo 2000, pois além da cozinha, também passam pelo refeitório e outros locais. Como resultado, eles

têm uma rotina mais complexa, com maior número de variáveis de tempo e movimento.

**Indivíduos do tipo 4000 (limpeza tradicional)** - A rotina desses indivíduos é mais simples, pois eles geralmente ficam apenas no final do dia para realizar a limpeza. Eles frequentam menos espaços e têm uma rotina mais curta, com menos variáveis associadas.

**Indivíduos do tipo 5000 (limpeza durante o dia)** - Esses indivíduos são aqueles realizam a limpeza ao longo do dia, circulando por diferentes espaços, especialmente no refeitório. A rotina deles é mais dinâmica, pois eles devem se adaptar rapidamente às mudanças, seja por emergências, doenças ou férias.

Para estes indivíduos cabe observar que é criado um *Dataframe* vazio. No caso do tipo 2000, por exemplo, como são 15 indivíduos, é feito um *for* 15 vezes para gerar esses indivíduos, com a criação de um código para cada um. Se o tipo do indivíduo é 2000, então o código do primeiro indivíduo deste tipo é 2001, o segundo 2002 e assim sucessivamente até 2016, de forma que são criados 15 indivíduos desse tipo.

Basicamente o comportamento de rotina desses indivíduos são variáveis que são criadas na base de dados. Cria-se uma variável chamada “hora de saída 1”, que é a hora que o indivíduo saiu do espaço 1, “velocidade de saída 1”, que é a velocidade que o indivíduo saiu do espaço 1, “nome do espaço 1”, que é o nome do espaço 1 que o indivíduo esteve, “código do espaço 1” e, assim, sucessivamente para cada um.

Importante observar que, dependendo do tipo de indivíduo, haverá mais ou menos variáveis, pois as movimentações entre espaços variam.

## **5.7 Consolidação da base de dados**

Após a criação e ajustes das rotinas de todos os tipos de indivíduos (comensais, cozinheiros, pessoal da limpeza), foi necessário unir todos os *DataFrames* que representam essas diferentes categorias de indivíduos em um único *DataFrame* final.

Isso foi feito utilizando a concatenação dos *DataFrames*, de forma que todas as variáveis (horários, velocidades, locais frequentados etc.) fossem combinadas corretamente. Após a concatenação, descartou-se quaisquer dados que não seriam utilizados no modelo final, como os dados temporários usados apenas para o sorteio de emergências, doenças ou férias.

Com isso obtém-se uma base de dados completa, que representa a rotina e os comportamentos dos indivíduos durante os 4 (quatro) anos modelados, considerando todas as variações de emergência, doença e férias. Esse é o modelo de simulação que pode ser utilizado para análise ou para prever cenários.

## **6. RESULTADOS**

Após rodar o modelo estruturado para este projeto chegou-se aos resultados, que serão comentados a seguir.

### **6.1 Contextualização da Base de Dados**

A base de dados utilizada neste estudo é composta por registros diários de indivíduos. Cada linha da base representa um dia de observação de um determinado indivíduo, contendo informações como:

- Horário de entrada e saída do primeiro local frequentado;
- Velocidade de entrada e saída;
- Indicação se houve ou não uma situação de emergência no dia.

A base possui 1014 dias de observação para cada indivíduo. Os indivíduos estão distribuídos em diferentes categorias, como já descritos no capítulo 3. Assim, o conjunto total de dados é extenso, com um grande volume de linhas representando diferentes indivíduos e categorias.

### **6.2 Divisão da Amostra**

Para o processo de treinamento e validação dos modelos, a amostra foi dividida em cinco grupos iguais, cada um contendo aproximadamente 20% das linhas da base. A divisão foi realizada de forma aleatória, sem reposição, garantindo que cada linha fosse atribuída a apenas um grupo.

A estratégia de validação consiste em utilizar quatro grupos (80% dos dados) para o treinamento do modelo e um grupo (20% dos dados) para o teste. Este processo foi repetido cinco vezes, alternando o grupo utilizado para teste, permitindo que cada grupo seja usado como conjunto de teste exatamente uma vez. Ao final, foi calculada a média das acurácias obtidas nos cinco testes, proporcionando uma avaliação robusta do desempenho dos modelos.

### 6.3 Modelos de Previsão

Os modelos utilizados incluem:

- Modelo de Probabilidade Linear (MPL);
- Modelo Probit;
- Modelo Logit.

Os modelos têm como variáveis explicativas os dados comportamentais dos indivíduos (horários de entrada e saída, velocidades, entre outros) e como variável dependente a indicação binária de emergência (1 para emergência, 0 para não emergência).

### 6.4 Procedimento de Treinamento e Teste

**Etapa de Treinamento** - Durante a etapa de treinamento, os quatro grupos selecionados são utilizados para estimar os parâmetros dos modelos. Como a variável dependente é conhecida, é possível estimar relações entre os comportamentos observados e a presença de emergência, permitindo que o modelo aprenda a identificar padrões associados a situações emergenciais.

**Etapa de Teste** - Na etapa de teste, o grupo restante é utilizado para avaliar a capacidade preditiva do modelo. As previsões são feitas com base nos parâmetros estimados, e os resultados previstos (emergência ou não) são comparados com os resultados observados para calcular a acurácia do modelo.

**Ponto de Corte** - Os modelos geram probabilidades estimadas para a variável dependente. Para transformar essas probabilidades em previsões binárias (0 ou 1), foi adotado um ponto de corte de 50%:

- Probabilidade estimada  $> 50\%$ : classificado como emergência (1);
- Probabilidade estimada  $\leq 50\%$ : classificado como não emergência (0).

Observação Importante: No caso do modelo de probabilidade linear, as estimativas de probabilidade podem ocasionalmente exceder o intervalo de 0 a 1. Embora isso represente uma limitação técnica, o ponto de corte corrige esse problema na prática. Em contrapartida, os modelos Probit e Logit são parametrizados para gerar probabilidades sempre dentro do intervalo de 0 a 1.

## **6.5 Avaliação de Desempenho**

A acurácia de cada modelo é calculada para cada rodada de validação cruzada e, posteriormente, é feita a média das acurácias obtidas. Isso fornece uma medida consolidada do desempenho preditivo do modelo. De forma geral, os modelos utilizados apresentaram acurácias elevadas, com médias superiores a 96% nos casos analisados.

## **6.6 Conclusão**

Os resultados indicam que a metodologia adotada é eficaz para avaliar e comparar o desempenho de diferentes modelos preditivos no contexto de previsão de emergências. Além disso, a divisão em grupos e o uso da validação cruzada garantem maior robustez na análise, reduzindo a dependência de subdivisões específicas da amostra.

Os resultados são apresentados a seguir:

- Pelo método MPL:
  - Média de acurácia: 0.964497862096726
  - Desvio padrão de acurácia: 0.0041404734871738885
  
- Pelo método Probit:
  - Média de acurácia: 0.9646330283659917
  - Desvio padrão de acurácia: 0.0042653790910075545
  
- Pelo método Logit:
  - Média de acurácia: 0.9648789597073142
  - Desvio padrão de acurácia: 0.004243165667533134

## 7. CONSIDERAÇÕES FINAIS

Os resultados do experimento, como indicado no capítulo 6, foram interessantes, indicando que a metodologia adotada é eficaz na previsão de emergências. O presente estudo teve a intenção de avaliar o potencial de desenvolvimento de uma Inteligência Artificial (IA) que colabore em processos de tomada de decisão, jogando sua luz sobre as possibilidades deste modelo e sua aplicação em situações de emergência, em especial em ambiente militar. Há uma gama de outras abordagens que podem ser consideradas e que devem ser objeto de estudo mais aprofundado.

A Inteligência Artificial aqui desenvolvida foi utilizada como um elemento para tomada de decisão, o que é um elemento importante a ser discutido. Assim, cabe refletir sobre um caso ocorrido em junho de 1982 durante a Guerra das Malvinas e a possibilidade de uma IA que tomasse decisões de forma imediata ao invés de subsidiar informações para o alto escalão militar.

Em um dado momento do conflito, um avião de combate britânico entrou no espaço aéreo brasileiro. Neste caso, se a tomada de decisão sobre abater o vetor aéreo fosse tomada exclusivamente por uma IA, uma possibilidade concreta seria o abate da aeronave, o que geraria um possível conflito entre os dois países. Fica a pergunta: Abater o avião Vulcan britânico seria o melhor para o Brasil naquele episódio?

O uso da IA na ponta final do processo, ou seja, como o efetivo tomador de decisão, pode ser um risco grande para determinadas ações. Este modelo buscou observar um ambiente e trazer elementos para assessorar os tomadores de decisão sobre respostas aos eventos observados, mas não uma ação deliberada ao final do processo.

Outro ponto que cabe reflexão é o quanto tal assessoramento prestado por uma IA seria capaz de influenciar os tomadores de decisão a uma determinada ação? O algoritmo pode chegar ao ponto de ser capaz de não somente de analisar qual seria a melhor escolha a se tomar, mas também de estudar o perfil do decisor para induzi-lo a uma ou outra decisão que a IA considere a melhor? São perguntas que são importantes serem discutidas ao longo deste percurso.

## REFERÊNCIAS

Gerais:

Al-amri, R., Murugesan, R.K., Man, M.; Abdulateef, A.F., Al-Sharafi, M.A.; Alkahtani, A.A. A. (2021). Review of Machine Learning and Deep Learning Techniques for Anomaly Detection in IoT Data. *Appl.Sci.* 2021, 11, 5320. <https://doi.org/10.3390/app11125320>.

Andrade, E. Fisher, B. & Blunsden. (2006). Detection of Emergency Events in Crowded Scenes. in *Crime and Security. The Institution of Engineering and Technology Conference on* . pp. 528-533.

Bernardi, S., Javierre, R. & Merseguer, J. (2023). tegdet: An extensible Python library for anomaly detection using time evolving graphs. *SoftwareX.* 22. 101363. [10.1016/j.softx.2023.101363](https://doi.org/10.1016/j.softx.2023.101363).

Burkov, A. (2019). *The hundred-page machine learning book* (Vol. 1, p. 32). Quebec City, QC, Canada: Andriy Burkov.

Chaib, C. J. A. (2015). A Defesa Nuclear, Biológica, Química e Radiológica nas instalações sensíveis da Marinha do Brasil. *Revista Âncoras e Fuzis – Corpo de Fuzileiros Navais Ano XIV – Número 46 – 2015*ISSN 2177-7608. pp 49-51.

Evans, M. J., & Rosenthal, J. S. (2004). *Probability and statistics: The science of uncertainty*. Macmillan.

Ibrahim, M., Alsheikh, A., Awaysheh, F.M., Alshehri, M.D. (2022). Solar Power Plants Anomaly Detection Using Machine Learning. *Energies* 2022, 15, 1082. <https://doi.org/10.3390/en15031082>.

Izbicki, R., & dos Santos, T. M. (2018). *Machine Learning sob a ótica estatística*. Ufscar/Insper.

Kang, B. & Choo, H. (2016). A deep-learning-based emergency alert system. Department of Computer Science and Engineering, Sungkyunkwan University, Suwon, 440-746, Republic of Korea.

Lazzarato, A., Boghossian, B. A. & Velastin, S. A. (1999). Detection of potentially dangerous situations involving crowds using image processing. Department of Electronic Engineering, King's College London (KCL, University of London).

- Lopez-Fuentes, L., Van de Weijer, J., Gonzalez-Hidalgo, M., Skinnemoen, H. & Bagdanov, A. D. (2017). Review on Computer Vision Techniques in Emergency Situations. *Multimedia Tools and Applications* (Researchgate).
- Marinha do Brasil (2020). CGCFN 10.3 - Manual de Defesa Nuclear, Biológica, Química e Radiológica (1ª Edição ed.). Rio de Janeiro: Comando-Geral do Corpo de Fuzileiros Navais.
- Patriota, M. P. (2022). Breve histórico da Defesa Nuclear, Biológica, Química e Radiológica na Marinha do Brasil. *Defesa NBQR em revista*. Ano II –Número 2 – 2022. pp 5-30.
- Rocha, S. F. (2008). Acidente radioativo com o Césio137: a participação da Marinha no atendimento às vítimas. *Navigator (Especial)*, 9-78
- Sun Tzu. *A arte da guerra*. Tradução de André Bueno. São Paulo: Editora Record, 2007
- Yilmaz, S. & Kozat, S. (2020). PySAD: A Streaming Anomaly Detection Framework in Python (Researchgate).

#### Referências MPL, Logit e Probit:

- Amemiya, T. (1981). "Qualitative response models: A survey." *Journal of Economic*
- Greene, W. H. (2020). *Econometric Analysis*. Pearson.
- Horowitz, J. L., & Savin, N. E. (2001). "Binary response models: Logit, probit, and semiparametric alternatives." *Journal of Economic Perspectives*, 15(4), 43–56.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression*. Wiley.
- Literature*, 19(4), 1483–1536.
- McFadden, D. (1974). "Conditional logit analysis of qualitative choice behavior." *Frontiers in Econometrics*.
- Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics*. Cambridge University Press.
- Wooldridge, J. M. (2019). *Introductory Econometrics: A Modern Approach*. Cengage.