

**UNIVERSIDADE DE PERNAMBUCO
ESCOLA POLITÉCNICA DE PERNAMBUCO
COORDENAÇÃO DE ENGENHARIA DE COMPUTAÇÃO**

CAIO EMANNUEL LIRA DE SANTANA

**UTILIZAÇÃO DE REDES NEURAS ARTIFICIAIS PARA A PREVENÇÃO DE
FRAUDES EM BENEFÍCIOS DA PREVIDÊNCIA SOCIAL**

RECIFE

2018

CAIO EMANNUEL LIRA DE SANTANA

**UTILIZAÇÃO DE REDES NEURAIS ARTIFICIAIS PARA A PREVENÇÃO DE
FRAUDES EM BENEFÍCIOS DA PREVIDÊNCIA SOCIAL**

Monografia apresentada ao Curso de Engenharia de Computação da Escola Politécnica da Universidade de Pernambuco, como requisito parcial para a obtenção do título de Bacharel em Engenharia de Computação.

**Orientador: Prof. Dr. Mêuser Valença
Coordenador: Prof. Dr. Daniel Ribeiro**

RECIFE

2018

MONOGRAFIA DE FINAL DE CURSO

Avaliação Final (para o presidente da banca)*

No dia 26 de dezembro de 2018, às 8:00 horas, reuniu-se para deliberar a defesa da monografia de conclusão de curso do discente **CAIO EMANNUEL LIRA DE SANTANA**, orientado pelo professor **Mêuser Jorge Silva Valença**, sob título **UTILIZAÇÃO DE REDES NEURAIS PARA A PREVENÇÃO DE FRAUDES NA PREVIDÊNCIA SOCIAL**. a banca composta pelos professores:

Sérgio Mário Lins Galdino

Mêuser Jorge Silva Valença

Após a apresentação da monografia e discussão entre os membros da Banca, a mesma foi considerada:

Aprovada



Aprovada com Restrições*

Reprovada

e foi-lhe atribuída nota: 9,0 (*naal*)

*(Obrigatório o preenchimento do campo abaixo com comentários para o autor)

O discente terá 7 dias para entrega da versão final da monografia a contar da data deste documento.

SÉRGIO MÁRIO LINS GALDINO

MÊUSER JORGE SILVA VALENÇA

* Este documento deverá ser encadernado juntamente com a monografia em versão final.

A Deus, a minha esposa Camila,
Aos meus pais, Severino Santana e Maria Suely de Santana, a eles todos os
créditos.

Dedico

AGRADECIMENTOS

A Deus, por iluminar o meu caminho e me permitir atingir mais esse objetivo me concedendo energia para enfrentar desafios diariamente.

A minha família, em especial meus pais, irmãos, sobrinhos pela atenção e por me incentivarem a realizar mais esse sonho.

A minha esposa Camila, grande companheira de luta, por caminhar ao meu lado nesta longa jornada.

Aos meus professores, por me oportunizarem expandir meus horizontes, compartilhando conhecimentos e experiências que me moldaram como profissional.

“Combati o bom combate, acabei a corrida e guardei a fé. Desde agora, a coroa da justiça me está guardada, a qual o Senhor, justo juiz, me dará naquele Dia; e não somente a mim, mas a todos que amarem a sua vinda.”

2 Timóteo 4:7-8.

SUMÁRIO

LISTA DE ILUSTRAÇÕES	VII
RESUMO.....	VIII
ABSTRACT	IX
1 INTRODUÇÃO	10
1.1 MOTIVAÇÃO.....	10
1.2 OBJETIVOS	14
1.3 ESTRUTURA DA MONOGRAFIA.....	14
2 FUNDAMENTAÇÃO TEÓRICA	16
2.1 BENEFÍCIOS PREVIDENCIÁRIOS.....	16
2.2 BENEFÍCIOS ASSISTENCIAIS.....	19
2.3 INSS DIGITAL - UMA NOVA FORMA DE ATENDER.....	20
2.4 REDES NEURAIS ARTIFICIAIS - RNA.....	28
3 METODOLOGIA.....	36
3.1 BASE DE DADOS	36
3.2 PROCESSAMENTO DE DADOS.....	39
3.3 ARQUITETURA MLP NA IDENTIFICAÇÃO DE FRAUDES EM REQUERIMENTOS DE BENEFÍCIOS PREVIDENCIÁRIOS	43
4 RESULTADOS	46
4.1 ANÁLISE DE REQUERIMENTOS DE BENEFÍCIOS ASSISTENCIAIS.....	46
4.2 ANÁLISE DE REQUERIMENTOS DE BENEFÍCIOS PREVIDENCIÁRIOS....	47
5 CONSIDERAÇÕES FINAIS	49
5.1 CONCLUSÕES	49
5.2 TRABALHOS FUTUROS	50
REFERÊNCIAS.....	51

LISTA DE ILUSTRAÇÕES

FIGURA 1 - DÉFICIT DA PREVIDÊNCIA SOCIAL NO ANO DE 2017	11
FIGURA 2 - TELA INICIAL DO GERENCIADOR DE TAREFAS.....	24
FIGURA 3 - TELA INICIAL DO MEU INSS.....	25
FIGURA 4 - NEURÔNIO BIOLÓGICO	29
FIGURA 5 - NEURÔNIO ARTIFICIAL.....	30
FIGURA 6 - EXEMPLO DE UMA ARQUITETURA MLP	32
FIGURA 7 - FASES DO ALGORITMO BACKPROPAGATION	33
FIGURA 8 - ARQUIVO DE TEXTO COM OS DADOS DE BENEFÍCIOS ASSISTENCIAIS A SEREM INSERIDOS NO WEKA.....	44
TABELA 1 - CONVERSÃO DE DADOS DA VARIÁVEL GRUPO FAMILIAR.....	40
TABELA 2 - CONVERSÃO DE DADOS DA VARIÁVEL RENDA PER CAPITA.....	41
TABELA 3 - CONVERSÃO DE DADOS DA VARIÁVEL ACT.	41
TABELA 4 - CONVERSÃO DE DADOS DA VARIÁVEL ESPÉCIE.....	42

RESUMO

UTILIZAÇÃO DE REDES NEURAS PARA A PREVENÇÃO DE FRAUDES NOS BENEFÍCIOS DA PREVIDÊNCIA SOCIAL.

Objetivo: Treinar uma Rede Neural Artificial para que a mesma seja capaz de, dado um conjunto de características, definir se um requerimento protocolado junto ao INSS representa uma fraude em potencial contra a Previdência Social brasileira.

Metodologia: Serão utilizados dois conjuntos de dados, um representando características de risco em requerimentos de benefícios assistenciais, e outro representando características de risco em benefícios previdenciários. Cada um destes conjuntos possui 60 amostras retiradas dos acervos da própria instituição. Destas 60 amostras, 30 serão utilizadas para treinar a rede, e 30 serão utilizadas para a validação do treinamento.

Resultados: Para o primeiro grupo de características, os valores de erro médio absoluto e erro médio quadrático foram de respectivamente, 0,184% e 0,434%. Das 30 amostras utilizadas para a validação, 24 foram classificadas corretamente e 6 foram classificadas incorretamente. Já no segundo grupo de benefícios, os valores de erro médio absoluto e erro médio quadrático foram de 0,16% e 0,39%, ao passo que, das 30 amostras, 26 foram classificadas corretamente e 6 incorretamente.

Conclusão: A arquitetura MLP se mostrou bastante útil no combate às fraudes da Previdência Social, sendo, necessário realizar o estudo do comportamento de outras arquiteturas de Redes Neurais, diante do problema exposto neste trabalho. Também é necessário um estudo visando maior aperfeiçoamento da arquitetura MLP, bem como dos dados apresentados como entrada, visando uma maior precisão da rede MLP para a solução do problema.

Descritores: Previdência Social, Redes Neurais Artificiais, Fraudes contra a previdência, Inteligência Artificial, Benefícios Previdenciários, Benefícios Assistenciais, INSS Digital, MLP.

ABSTRACT

USE OF ARTIFICIAL NEURAL NETWORKS TO PREVENT FRAUDS IN BENEFITS OF THE THE SOCIAL INSURANCE

Purpose: Prepare an Artificial Neural Network to be able to discovery if a request made against the Brazilian social insurance represents a risk of being a fraud, according to a group of features presented to the Neural Network previously.

Method: Two sets of data will be used, one representing risk characteristics in welfare benefits requirements, and another representing risk characteristics in social security benefits. Each of these sets has 60 samples taken from the collections of the institution itself. Of these 60 samples, 30 will be used to train the network, and 30 will be used for the validation of the training.

Results: For the first group of characteristics, the values of Mean Absolute Error and Root Mean Squared Error were, respectively, 0.184% and 0.434%. Of the 30 samples used for validation, 24 were classified correctly and 6 were classified incorrectly. In the second group of benefits, the values of mean absolute error and mean square error were 0.16% and 0.39%, whereas, of the 30 samples, 26 were classified correctly and 6 incorrectly

Conclusion: The MLP architecture proved to be very useful in combating Social Security fraud, but it is necessary to study the behavior of other Neural Networks architectures, in view of the problem presented in this paper. It is also necessary to further improve the MLP architecture, as well as the data presented as input, aiming at a greater accuracy of the MLP network to solve the problem.

Key words: Social Insurance, Artificial Neural Networks, Fraud against Social Insurance, Artificial Intelligence, Social Insurance Benefits, Assistencial Benefits, INSS Digital, MLP

1 INTRODUÇÃO

O presente capítulo traz uma explanação acerca dos fatores que motivaram a realização deste trabalho. Em seguida, são explanados os objetivos a serem atingidos, bem como a disposição dos capítulos desta monografia

1.1 MOTIVAÇÃO

Um dos principais pontos a ser abordado pelo governo brasileiro para reduzir o déficit nas contas públicas é a redução de gastos com a Previdência Social. Segundo dados apresentados pela Secretaria da Previdência, só em 2017, foi registrado um déficit de aproximadamente R\$ 182 bilhões. Esse “rombo” nas contas da seguridade social é 21,8% maior que o registrado no final do ano de 2016 [1].

O dado estarrecedor apresentado pelo Tribunal de Contas da União (TCU), estimando que cerca de R\$ 56 bilhões tenham sido gastos em decorrência de fraudes e erros nos procedimentos administrativos ofertados à população [2]. O montante apresentado pelo TCU representa algo em torno de 30% do prejuízo identificado ao final de 2017, o que agrava, ainda mais a situação fiscal na qual se encontra o Brasil.

Segundo o TCU, as fraudes envolvem diversas situações onde organizações criminosas envolvendo servidores públicos e pessoas ligadas a estes, os chamados intermediários ou atravessadores, buscam junto a população, voluntários, com o intuito de utilizá-los em práticas delituosas contra a Previdência Social. Ricos simulando pobreza, presidiários reconhecendo a paternidade de crianças, ciganos que vagam pelo país simulando situações de falecimento e nascimento, são algumas das situações que levam à ocorrência de fraudes e erros na prestação de serviços relacionados à seguridade social do Brasil.

Insta salientar, ainda, a forte característica burocrática da administração pública brasileira, que muitas vezes acaba dificultando a identificação e prevenção de ações orquestradas por verdadeiras quadrilhas, que acabam por prejudicar a prestação de serviços essenciais à nação. Arquivos perdidos ou propositalmente destruídos, a emissão de documentos de identificação falsificados e a falta de compartilhamento de informações entre os setores governamentais fazem com que

os grupos criminosos atuem, muitas vezes, sem ser detectados pelos sistemas de controle existentes.

FIGURA 1 - DÉFICIT DA PREVIDÊNCIA SOCIAL NO ANO DE 2017.



FONTE: REPRODUZIDO DE [2].

Com isso, o Instituto Nacional do Seguro Social (INSS) vêm implementando medidas de combate a operações criminosas realizadas no âmbito de seus serviços. Em parceria com o TCU e com a Polícia Federal, o instituto tem buscado investigar as fraudes e punir os envolvidos, visando sanar cada vez mais o saldo negativo bilionário em suas contas.

Assim, com o intuito de otimizar seus fluxos de atendimento, bem como diminuir a ocorrência de fraudes e erros em seus processos, o Instituto desenvolveu o projeto INSS Digital – Uma nova forma de atender. A implantação do projeto INSS Digital tem se mostrado um verdadeiro ponto de inflexão na forma de atendimento dessa autarquia. O projeto visa o desenvolvimento de uma nova forma de atendimento, visando aumentar a capacidade do órgão de analisar e reconhecer direitos [3].

O projeto prevê, ainda, um maior controle sobre seus procedimentos, já que toda a documentação referente ao reconhecimento de direitos estará disponível em uma plataforma on-line, para acesso e consulta pela população e pelos órgãos de controle. O INSS Digital reforça, também, a impessoalidade do servidor público na análise dos processos, visto que este receberá uma demanda que poderá advir de locais diversos de sua lotação original, sem que este tenha alguma margem de preferência sobre qual pessoa será ir a atender.

O funcionamento do INSS Digital baseia-se em uma distribuição de carga de trabalho entre os funcionários de todo o Brasil. Tal distribuição será feita de modo a equalizar a capacidade de trabalho das diversas localidades do país, já que em alguns lugares existe uma sobrecarga de demandas, enquanto que em outros, funcionários ficam ociosos, devido à baixa carga de trabalho.

Apesar de representar uma significativa inovação no serviço público, o INSS Digital não está totalmente imune à ação de fraudadores. Um dos pontos sensíveis, e que será abordado nesta Monografia, está intimamente relacionado com a alta capilaridade do INSS. Segundo dados do site oficial da autarquia, são mais de 1.500 pontos de atendimento espalhados por todo o Brasil [4], tendo estes a atribuição de realizar todos os atendimentos que são ofertados à sociedade. Essa alta capilaridade, contribui para a formação de quadrilhas com atuação, muitas vezes, voltadas à jurisdição de um município ou de um grupo de municípios.

Essa atuação regionalizada das quadrilhas, acaba por diversificar o *modus operandi* dos fraudadores que, de acordo com as peculiaridades de cada lugar, estudam a melhor forma de burlar as regras do sistema. Tomemos o exemplo do estado do Amazonas, onde índios recém-nascidos são adotados exclusivamente com o intuito de que mulheres, agindo de má-fé, pleiteem benefícios como Salário-Maternidade. Há ainda, casos como as quadrilhas de ciganos que atuam sobretudo no estado do Maranhão, falsificando documentos e recrutando pessoas para atuarem como estelionatários junto à Previdência Social.

Cabe destacar, ainda, que a forma como as organizações criminosas agem também variam de acordo com a quantidade de servidores públicos que possam estar envolvidos no esquema malicioso. Em alguns lugares, como na cidade do Cabo de Santo Agostinho, em Pernambuco, a Polícia Federal, através das operações “*Manager*” e “*Insistência*” [5][6], verificou a atuação de membros de Sindicato dos Trabalhadores Rurais, juntamente com funcionários do INSS, na concessão de aposentadorias para falsos trabalhadores rurais.

Já com a operação “*Citrus Prev*” [7] realizada na cidade de Limoeiro, no agreste Pernambucano, a Polícia Federal identificou a presença de funcionários terceirizados do INSS, atuando junto a intermediários e servidores da previdência, na concessão de benefícios assistenciais a falsos deficientes. No estado de São Paulo, a inteligência previdenciária, através da operação “*Pseudea*” [8], desarticulou

um grupo criminoso que fazia o uso de dublês, pessoas que portavam documentos falsificados e fingiam estar incapacitadas ou mentalmente transtornadas, inclusive com membros engessados, para ludibriar os médicos peritos do INSS.

Pode-se observar, portanto, que os criminosos buscam, a todo momento, atualizar e adaptar suas técnicas de violação dos sistemas, de acordo com as condições e características de cada região. Esse comportamento mutante tem grande potencial para ocasionar sérios danos aos novos fluxos estabelecidos pelo INSS Digital, já que, com a distribuição da demanda a nível nacional, um servidor lotado em uma cidade do Rio Grande do Sul, poderá acabar analisando um ou vários requerimentos fraudulentos advindos do Pará, por exemplo. Sem saber como operam os fraudadores paraenses, o funcionário gaúcho poderá ser conduzido ao erro, concedendo um benefício a alguém que pode estar mal-intencionado, operando em desacordo com a lei.

Com isso, esta monografia apresenta uma alternativa para que, utilizando Redes Neurais, do tipo *Multi-Layer Perceptron* (MLP), crie-se um sistema de alerta de irregularidades. Esse sistema tomará como entrada, dados relativos aos requerimentos de benefícios apresentados pela população e emitirá, com base nas experiências de fraudes identificadas pela inteligência previdenciária, alertas, classificados em níveis preestabelecidos, de acordo com as características mais presentes na forma de atuação das quadrilhas.

Esse sistema deverá, utilizando conceitos de Inteligência Artificial [IA] apresenta para os servidores responsáveis pelo requerimento qual o grau de risco existente naquele pedido. Com isso, caso um pedido seja realizado por uma mulher que alega ser mãe adotiva de um bebê indígena nascido em uma humilde comunidade ribeirinha do Amazonas, poderá ser enquadrado como um risco de fraude. Ademais, uma aposentadoria rural solicitada por cidadão residente no Cabo de Santo Agostinho e vinculado ao Sindicato dos Trabalhadores Rurais do referido município, também poderá ser classificado como risco de fraude, a depender de outras características a serem analisadas pelo sistema. Para tanto, serão utilizadas informações extraídas de operações realizadas pela inteligência previdenciária e pela Polícia Federal, para se montar um conjunto de características presentes nas ações dos fraudadores. Insta salientar que não será necessário o uso de dados

peçoais da população, apenas as informações públicas amplamente disponíveis nos meios de comunicação.

1.2 OBJETIVOS

1.2.1 OBJETIVO GERAL

Este projeto tem como objetivo preparar uma rede neural artificial para que, dada uma série de parâmetros, que representam características recorrentes em fraudes contra a previdência social, a mesma possa identificar a probabilidade de ocorrência de fraude em requerimentos de benefícios previdenciários e assistenciais, de acordo com as características de entrada.

1.2.2 OBJETIVOS ESPECÍFICOS

1. Definir uma série de parâmetros que deve ser considerada como sinais de entrada da rede neural artificial.
2. Promover o aprendizado do sistema com o intuito de encontrar os menores valores de erro possíveis.
3. Executar simulações com dados reais, em busca de se obter resultados satisfatórios.
4. Propor novos estudos que possam ser realizados com base nesse projeto, visando uma maior difusão do uso das técnicas de Inteligência Artificial no tratamento das demandas pleiteadas pela população junto ao INSS.

1.3 ESTRUTURA DA MONOGRAFIA

O capítulo 2 apresentará as definições necessárias para melhor compreensão dos experimentos a serem desenvolvidos, a saber: INSS Digital, benefícios previdenciários e Redes Neurais Artificiais, enfatizando na arquitetura MLP, escolhida para construção deste projeto. No capítulo 3 será descrita a metodologia utilizada com o intuito de atingir os objetivos estabelecidos na seção 1.2, além de explanar como será realizado o processamento de dados, a configuração e aplicação dos dados colhidos na arquitetura MLP. O capítulo 4

tratará dos resultados obtidos após as simulações, comparando os resultados obtidos com as condições reais, obtidas com base em levantamentos realizados por operações de combate às fraudes previdenciárias. E, por fim, o capítulo 5 trará as conclusões alcançadas e sugestões para desenvolvimento de trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Esse capítulo apresenta definições relevantes para a compreensão dos experimentos realizados neste trabalho. A seção 2.1 discorre sobre uma visão geral dos benefícios ofertados pelo Instituto Nacional do Seguro Social à sociedade, bem como sobre quais os requisitos exigidos para que se tenha acesso a algum desses benefícios. A seção 2.2 apresenta definições sobre o projeto INSS Digital – Uma nova forma de atender, explorando as mudanças trazidas com os novos fluxos e tratando de conceitos inerentes ao referido projeto. A seção 2.3 traz os conceitos referentes às Redes Neurais Artificiais, abordado de forma especial a arquitetura MLP, bem como suas formas de treinamento.

2.1 BENEFÍCIOS PREVIDENCIÁRIOS

Os benefícios previdenciários são prestações financeiras pagas à população como resultado de um sistema contributivo, onde as pessoas destinam contribuições financeiras para a Previdência Social, com o intuito de fazer jus ao recebimento de prestações financeiras, com natureza salarial. De acordo com a Lei Ordinária nº 8.213 de 24 de julho de 1991 [9], os benefícios previdenciários são os seguintes:

1. Aposentadoria por Idade.
2. Aposentadoria por Tempo de Contribuição.
3. Salário Maternidade.
4. Pensão por Morte.
5. Auxílio-Doença.
6. Aposentadoria por Invalidez.
7. Salário-Família.
8. Auxílio-Reclusão.
9. Auxílio-Acidente.
10. Aposentadoria Especial

Insta salientar que toda essa gama de benefícios fazem parte do Regime Geral de Previdência Social (RGPS), regime criado pela Constituição Federal de

1988, a qual determina que o sistema se baseará na formação de reservas para garantir os benefícios citados [10].

A Aposentadoria por Idade é um direito reservado às pessoas que contribuíram para o RGPS por um período mínimo de 180 meses e que possuam 65 anos, no caso de homens, ou 60 anos no caso de mulheres. O benefício também pode ser concedido aos agricultores que comprovarem desempenho de atividade rural durante o período mínimo de 180 meses.

Já a Aposentadoria por Tempo de Contribuição é uma benesse previdenciária concedida em favor daqueles que possuem o tempo mínimo de 35 anos, se homem, ou 30 anos, se mulher, de contribuição. Insta salientar que a Aposentadoria por Tempo de Contribuição não exige a condição de idade mínima. O tempo mínimo de contribuição necessário, poderá, ainda, ser reduzido em virtude do desempenho de atividades insalubres, perigosas ou penosas.

O Salário-Maternidade é o benefício concedido em favor das mulheres que contribuam para o RGPS em razão do nascimento do seu filho, ou da adoção de uma criança. Segundo o Professor Frederico Amado [11], o Salário-Maternidade, concedido, em regra, por um período de 120 dias, restitui a mulher durante o período em que ela deve voltar a sua atenção ao infante, sendo presumida a sua condição de incapacidade para o trabalho. Excepcionalmente, o salário-maternidade poderá ser concedido aos pais da criança, em caso de falecimento da mãe no momento do parto, por exemplo.

Pensão por morte é o nome dado à prestação pecuniária paga pelo RGPS nos casos de falecimento de pessoa sob a qual viviam às expensas os filhos menores de 21 anos, o cônjuge ou companheiro, os pais ou os irmãos. Cabe destacar que o cônjuge/companheiro, bem como os filhos menores de 21 anos possuem dependência econômica presumida [11], ou seja, não necessitam comprovar que dependiam financeiramente do falecido para fazer jus à pensão. Já nos casos de pais ou irmãos, estes precisaram comprovar que dependiam economicamente do falecido.

O Auxílio-Doença e a Aposentadoria por Invalidez são benefícios concedidos àqueles que, por motivo de doença ou acidente de trabalho, ficarem incapacitados para o desempenho de suas atividades laborais. No caso específico da Aposentadoria por Invalidez, a incapacidade citada deverá ser permanente,

impedindo que o profissional volte a desempenhar as suas atividades. Ambos os benefícios necessitam de perícia médica prévia, a ser realizado por médico perito pertencente aos quadros do INSS.

O Salário-Família, por sua vez, é concedido como uma forma de complementação para cobrir despesas domésticas com os filhos menores de 14 anos ou, se inválidos, de qualquer idade [11]. Para tanto, o requerente deverá ser enquadrado como de baixa-renda. O salário-família não tem natureza de remuneração, sendo apenas um adicional para auxiliar os trabalhadores menos favorecidos a manter seus filhos menores devidamente matriculados em instituições de ensino.

Já o Auxílio-Reclusão, consiste em benefício previdenciário pago em favor dos dependentes de trabalhador recolhido à prisão, desde que o mesmo não receba remuneração, aposentadoria ou auxílio-doença [11]. Aqui também deverá ser verificada a condição de baixa renda do preso, conforme os limites legais estabelecidos na legislação previdenciária [9]. O auxílio-reclusão segue o mesmo raciocínio da pensão por morte, visto que aquele é pago em favor dos dependentes do recluso, a saber: cônjuge/companheiro, filhos menores de 21 anos, pais e irmãos, respeitando a mesma regra de comprovação de dependência econômica com relação aos pais e irmãos, já abordada na discussão sobre a pensão por morte.

O Auxílio-Acidente é um benefício pago como uma indenização, em virtude de infortúnio redutor da capacidade laborativa do trabalhador, por motivo de acidente de qualquer natureza ou causa. Cabe destacar, ainda que o Auxílio-Acidente é a única benesse previdenciária para com natureza indenizatória [11], sendo pago em virtude da consolidação de lesões, na forma de sequelas, que implicam em uma redução da capacidade laborativa do trabalhador.

Por fim, temos a Aposentadoria Especial, benefício concedido em favor dos trabalhadores que desempenham as suas funções sob condições especiais, capazes de prejudicar a sua saúde ou integridade física, bem como nos casos de trabalhador portador de deficiência. Assim, caso haja exposição a agentes nocivos, ou de acordo com o grau de deficiência de um trabalhador, o tempo de contribuição exigido para fazer jus a uma aposentadoria especial poderá ser de 15, 20 ou 25 anos de contribuição para o RGPS.

2.2 BENEFÍCIOS ASSISTENCIAIS

Apesar do caráter contributivo ser uma das principais características do RGPS, onde as pessoas devem verter contribuições financeiras para o sistema, a fim de um dia poderem fruir prestações pecuniárias decorrentes do pagamento de benefícios previdenciários, há um conjunto específico de serviços prestados à população, no âmbito do RGPS, que independem da realização, por parte da população, de contribuições financeiras para o sistema.

Esse conjunto de serviços gratuitos faz parte da chamada Assistência Social, sistema que compreende ações relacionadas à defesa de pessoas consideradas vulneráveis, ou seja, que vivem à margem da sociedade, em situação calamitosa, necessitando, assim, de suporte governamental para que tenham um mínimo de dignidade em sua condição de vida.

De acordo com a Constituição Federal de 1988 [10], a assistência social deverá ser prestada em favor daqueles que dela necessitarem, independentemente da ocorrência de contribuições para a seguridade social, devendo os recursos desse amparo governamental serem obtidos através de recursos provenientes das contribuições para o RGPS, pagas pelos trabalhadores, acrescido de outras fontes de custeio.

A teia assistencialista do governo brasileiro reflete o caráter inclusivo presente no sistema de seguridade social do País. Pessoas inválidas, que nunca puderam desempenhar qualquer atividade laborativa, ou até mesmo idosos que não possuem nenhum recurso financeiro para cobrir suas despesas com alimentação e saúde, podem ter acesso, através de requerimento formalizado junto ao INSS, a uma prestação pecuniária mínima, para elevá-los a uma condição de vida mais justa e digna. Com isso, temos que a Assistência Social prevista em nossa Carta Magna, se coaduna com os princípios basilares do Estado brasileiro.

Segundo a Lei Orgânica da Assistência Social brasileira [11], os benefícios que configuram pagamento, pelo RGPS, de prestações pecuniárias permanentes, ofertados para aqueles que se encontram em situação de vulnerabilidade, são os seguintes:

1. Benefício de Prestação Continuada à pessoa com deficiência.
2. Benefício de Prestação Continuada à pessoa idosa.

O Benefício de Prestação Continuada (BPC) é configurado como uma prestação financeira de pagamento mensal, com valor correspondente a 1 (um) salário-mínimo, que será pago aos que não possuem meios de prover sua própria manutenção, ou de tê-la provida através de seus familiares [11].

O BPC à pessoa com deficiência será pago às pessoas que possuem impedimentos de longo prazo, podendo ser configurado através de deficiência física, mental, intelectual ou sensorial, o qual poderá dificultar a sua participação na sociedade nas mesmas condições das outras pessoas. Cabe salientar, ainda, que, por se tratar de amparo mantido pelo Estado, o BPC à pessoa com deficiência deverá levar em consideração a existência de rendas auferidas pelo grupo familiar daquele que pleitear o recebimento, não podendo a renda per capita de sua família ser igual ou superior a $\frac{1}{4}$ [um quarto] do salário-mínimo [11].

Nessa mesma esteira, o BPC à pessoa idosa será pago àqueles que possuam, no mínimo, 65 anos de idade, que não possuam meios para se sustentar. O BPC pago à pessoa idosa não leva em consideração a existência de qualquer tipo de deficiência, devendo observar apenas a idade mínima e as mesmas restrições relativas à renda do grupo familiar existentes para o BPC à pessoa com deficiência.

Cabe destacar, também, que nos casos de BPC à pessoa com deficiência, a condição incapacitante deverá ser investigada por avaliação médica e avaliação social, procedimentos realizados por médicos peritos e assistentes sociais vinculados ao INSS. Estes procedimentos visam identificar se há realmente a incapacidade e qual o grau de impedimento que ela acarreta para a vida do requerente [11].

2.3 INSS DIGITAL – UMA NOVA FORMA DE ATENDER

O Instituto Nacional do Seguro Social é considerado, atualmente, a maior autarquia do Poder Executivo Federal, com um quadro de aproximadamente 36.000 servidores ativos e uma rede de atendimento composta por mais de 1500 agências distribuídas por todo o Brasil. Segundo dados fornecidos pela própria instituição, só

em 2016 foram realizados cerca de 44 milhões de atendimentos presenciais, são pedidos de informações, realizações de perícias médicas, requerimentos de benefícios previdenciários e assistenciais e outros serviços fornecidos à sociedade.

Segundo levantamento realizado pela revista Valor Econômico [12], em 2017, os valores pagos pelo INSS a títulos de benefícios representaram mais de 25% do Produto Interno Bruto (PIB) em 500 cidades brasileiras. O levantamento comparou, ainda, o montante disponibilizado à população pelo INSS, com o Fundo de Participação dos Municípios (FPM), verba relacionada à transferência constitucional de recursos da União para os Estados e Distrito Federal, formada por cerca de 22% da arrecadação do Imposto de Renda (IR) e do Imposto sobre Produtos Industrializados (IPI) [10]. No comparativo, percebeu-se que, de 5.570 municípios, em 4.216, a receita proveniente dos benefícios previdenciários e assistenciais foi superior aos repasses do FPM em 2015.

Dada a magnitude dos números, é notória a pujança do INSS na economia nacional, tendo em vista seu poder de alavancar as receitas de muitos municípios do País, bem como a capilaridade no seu atendimento, já que o órgão possui pontos de atendimento espalhados por todo o Brasil. A grande demanda, à forte burocracia inerente ao serviço público em geral e a mecanização na realização dos processos, acaba por influenciar negativamente à prestação dos serviços da Previdência Social à sociedade.

Dados fornecidos pela instituição, demonstram que, em 2016, uma pessoa esperava mais de 60 dias em média, para ser atendida em um serviço agendado, em uma Agência da Previdência Social (APS). O cenário mostrou-se ainda mais preocupante, tendo em vista o envelhecimento do quadro de pessoal do órgão, que leva a uma previsão de que, a partir de janeiro de 2019, surja um movimento de aposentadoria em massa dos servidores, o que deve levar a um *déficit* de pessoal da ordem de 12 mil servidores, aproximadamente 1/3 (um terço) do quadro de pessoal existente.

Diante dessa situação, a Direção Central do INSS passou a discutir a criação de um projeto que fosse capaz de otimizar e automatizar os fluxos de atendimento, bem como aumentar a capacidade dos servidores para atender a grande demanda criada diariamente. Surgiu, assim, o projeto INSS Digital – Uma nova forma de atender. Segundo a própria instituição, o projeto tem como um dos

principais pilares a implementação de um processo eletrônico, onde toda a análise dos requerimentos de benefícios será feita pela Internet, sem a necessidade de comparecimento às agências.

O segundo pilar do projeto seria a distribuição de demandas entre as unidades. Essa ideia baseia-se no estudo de indicadores de produtividade e quantidade de demanda nas agências de todo o Brasil. Após análise destes parâmetros, o órgão identificou que havia uma distribuição bastante desigual de sua carga de trabalho, onde os gargalos estão situados nos grandes centros urbanos, como capitais e regiões metropolitanas, ao passo que, os servidores lotados em unidades mais interioranas, tendem a receber uma quantidade de trabalho bem menor, chegando, inclusive, a ficar ociosos em suas funções.

Assim, tomando por base os dois pilares centrais já citados, o INSS, em 2017, iniciou o processo de implantação dos novos fluxos de atendimento no Brasil. A experiência-piloto se deu nas unidades vinculadas à Gerência Executiva de Mossoró/RN, em janeiro de 2017 [13]. Nesse primeiro momento, ocorreram poucas modificações do atendimento nas agências. Com isso, o requerimento era solicitado, presencialmente, na própria APS, sendo a demanda deslocada, posteriormente, para um grupo de trabalho, montado na sede da Gerência Executiva já citada, em uma estrutura denominada “polo de análise”. Cabe destacar, que o requerimento, antes feito de forma totalmente manual, passou a consistir na digitalização da documentação do cidadão, anexando o arquivo digital no sistema responsável pelo gerenciamento das demandas criadas no novo formato digital.

Posteriormente, com a análise dos resultados obtidos na experiência-piloto, percebeu-se que o projeto de fato poderia ser expandido para todo o País. Iniciaram-se, com isso, os processos de “virada de chave”, ou seja, conversão do fluxo mecanizado, para o formato de atendimento digital. Diversas APS começaram a entrar no novo modelo, ademais, novos pólos de análise foram criados. Atualmente, quase que a totalidade das APS do Brasil aderiram ao novo formato de atendimento ao público.

Apesar de quase todas as APS do País já terem “virado a chave”, ou seja, já aderiram ao novo fluxo digital, discrepâncias ainda são encontradas. Existem algumas agências que, por estarem localizadas em cidades localizadas no interior remoto do país, não possuem recursos de internet suficiente para se tornar uma

APS digital. Outras, ainda, por questões de falta de mão-de-obra, entraram no novo modelo de forma apenas parcial, ou seja, realiza parte do atendimento nos moldes anteriores e outra parte já no fluxo atual. Ademais, os sistemas responsáveis por materializar o projeto INSS Digital, ainda estão passando por ajustes, o que acaba por gerar transtornos à população.

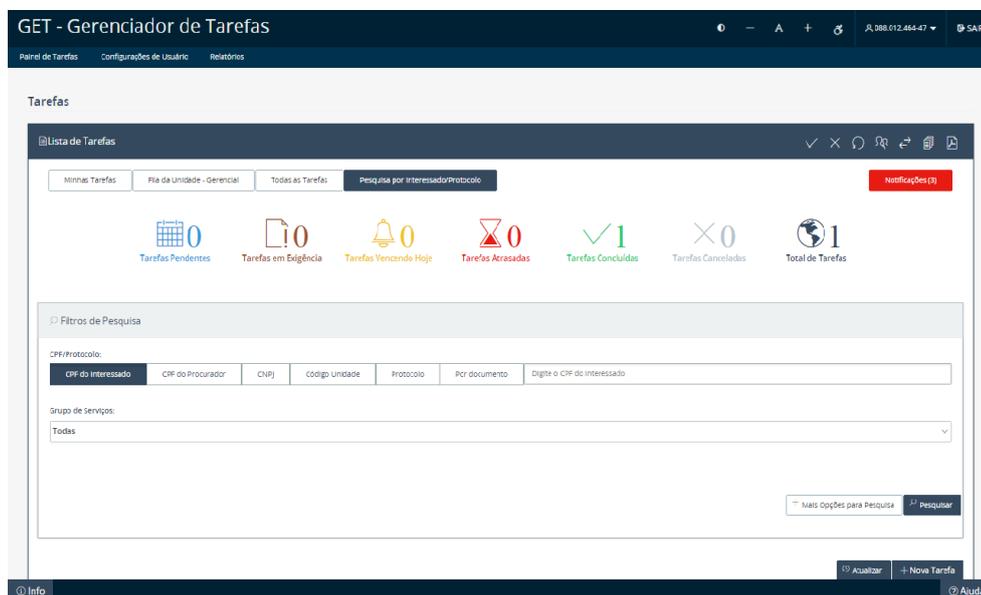
2.3.1 GERENCIADOR DE TAREFAS – GET

Nessa seção, é apresentada a plataforma fundamental para a materialização das mudanças promovidas pelo INSS Digital. É denominado o Gerenciador de Tarefas (GET). O Gerenciador consiste em uma ferramenta capaz de auxiliar os servidores e gestores a controlar suas demandas, exibindo, de forma bastante simples, toda a carga de trabalho que um funcionário possui no momento. Os gestores de unidade, ou gerentes de APS, possuem acesso privilegiado, pois conseguem visualizar toda a demanda existente em sua Agência.

O GET funciona dividindo a carga de trabalho em vários “status”, que variam de acordo com a situação em que cada requerimento se encontra. Funciona de forma semelhante aos gerenciadores de tarefas das metodologias ágeis de desenvolvimento de software. A seguir, eram apresentados todos os status onde um requerimento pode se encontrar. Para facilitar a compreensão dos status, é preciso entender o conceito de tarefas.

A tarefa é uma demanda criada quando do requerimento realizado por cidadão, via internet, ou em uma agência do INSS. As tarefas não compreendem apenas os requerimentos dos benefícios citados na seção 2.1, como também às solicitações de Revisão, Recurso contra decisão denegatória, Certidão de Tempo de Contribuição, Atualização de Dados Cadastrais, Perícias Médicas, entre outras. A ideia é a de que, quando da implantação completa do INSS Digital, todas as solicitações apresentadas pela sociedade junto ao INSS, sejam realizados via fluxo digital, ou seja, cada solicitação irá representar uma tarefa. No momento, essa associação de uma tarefa a qualquer solicitação ainda não está totalmente implementada, mas vem sendo feita de forma contínua.

FIGURA 2 - TELA INICIAL DO GERENCIADOR DE TAREFAS



FONTE: ACESSO CORPORATIVO DO INSS.

Os status nos quais uma tarefa poderá se encontrar são os seguintes:

1. Tarefa Pendente: Trata-se do estado inicial, onde se encontram todas as tarefas que já foram criadas, mas que, ainda não foram analisadas por nenhum funcionário do INSS. Nesse status, a tarefa poderá estar atribuída, ou não, a um funcionário, denominado responsável pela tarefa.
2. Tarefas Atrasadas: São todas aquelas tarefas que já ultrapassaram o prazo definido para a conclusão, sem que o servidor as tenha finalizado. Cabe destacar que, o prazo máximo para que o INSS conclua a análise de um requerimento é, em regra 45 dias. Esse prazo possui algumas exceções, como é o caso de requerimentos de salário-maternidade, que deverá ser concluído em 30 (trinta) dias. Importante destacar que, nesse status, as tarefas poderão, ou não possuírem responsáveis.
3. Tarefas Vencendo Hoje: São tarefas cujo prazo citado no item anterior se esvai no dia atual. No dia seguinte ao do vencimento do prazo, as tarefas serão levadas ao status Atrasada.

4. Tarefas Em Exigência: São aquelas que dependem de uma ação de pessoa diversa daquela que se encontra analisando a tarefa, para poder ser concluída. As exigências podem ser diversas, como a apresentação de documentação pelo requerente, ou a realização de pesquisas externas para que seja verificada uma condição alegada quando do requerimento. Uma vez suprida a exigência, a tarefa voltará para o status, pendente, atrasada ou vencendo hoje, a depender do seu prazo de conclusão.
5. Tarefas Concluídas: Representam todas as demandas que foram concluídas pelo servidor. Uma vez concluídas, as tarefas não mais podem ser modificadas ou reabertas.

2.3.2 MEU INSS

Seguindo a linha de mudanças de fluxos ocasionada pela implantação do INSS Digital, foi desenvolvida uma plataforma direcionada às pessoas, para que estas acessem e acompanhem todo o andamento de suas demandas, bem como obtenham informações relativas a todo o seu período laboral, sem que haja a necessidade de comparecimento da população a uma agência da previdência social.

FIGURA 3 - TELA INICIAL DO MEU INSS.



FONTE: <https://meu.inss.gov.br/>

O Meu INSS surgiu com a mesma ideia do Gerenciador de Tarefas, ou seja, criar uma ferramenta capaz de automatizar os processos e diminuir o tempo de

atendimento das demandas. O Meu INSS, no entanto, tem como um dos seus principais objetivos, proporcionar maior conforto para as pessoas, já que a plataforma disponibiliza vários serviços antes realizados apenas no âmbito de uma APS. A intenção é causar um “esvaziamento” das agências, para que a sociedade passe a fazer uso de uma ferramenta capaz de atender seus questionamentos com poucos cliques.

Para tanto, a plataforma é disponibilizada na internet, através do link <https://meu.inss.gov.br/> e do aplicativo Meu INSS, disponível nas principais lojas de aplicativos. Diz-se que o Meu INSS é fundamental para o bom funcionamento do INSS Digital como um todo, já que o mesmo tem promovido intensa diminuição na quantidade de atendimentos realizados nas APS de todo o Brasil, sendo um dos pilares dessa renovação de fluxos.

2.3.3 ACORDOS DE COOPERAÇÃO TÉCNICA – ACT

Conforme já foi explanado, o projeto INSS Digital tem como um de seus objetivos permitir que a população tenha acesso às suas informações, bem como protocolem seus requerimentos, sem a necessidade de comparecer a uma agência. Uma ferramenta criada pelo INSS para fortalecer ainda mais essa cultura de acesso aos serviços através dos canais remotos, o projeto lançou mão dos chamados Acordos de Cooperação Técnica (ACT).

A ideia central é de tais acordos sejam firmados junto a instituições da sociedade, delegando a esses entes, parte da execução dos serviços prestados nas agências do INSS. Assim, com a construção de um ACT, a população poderá requerer benefícios e serviços diretamente dessas instituições, evitando, assim, a fila eletrônica existente, já que, para serviços agendados, a sociedade chega a esperar mais de 6 meses entre o agendamento de um serviço e o seu dia de atendimento.

A intenção do INSS é firmar acordos com várias entidades, como prefeituras, empresas, sindicatos e a Ordem dos Advogados do Brasil. Uma vez constituído um acordo, essas entidades passam a fazer parte do fluxo de atendimento do INSS Digital. Esse acordo visa, ainda, dar uma maior capilaridade à prestação de serviços. A título de informação, o projeto piloto na Gerência-Executiva Mossoró/RN, já citado anteriormente, expandiu sua rede de atendimento de 18 agências, para mais de 48

locais de atendimento ao público [3]. Isso sem contar a quantidade de demandas solicitadas pelo próprio cidadão, através do Meu INSS.

Apesar de representar uma importante ferramenta na implementação dos novos fluxos de atendimento do INSS Digital, os ACTs possuem um risco associado, já que, a instituição “abre as portas” de seus sistemas corporativos para que terceiros venham a utilizá-los. Com isso, intermediários e pessoas má intencionadas poderão fazer uso das prerrogativas estabelecidas em um Acordo para pôr em execução práticas delituosas contra a Previdência Social.

Ademais, estes Acordos tendem a distanciar a população do atendimento presencial em uma Agência da Previdência Social, já que a população poderá registrar suas solicitações através de diversos pontos de atendimento espalhados pelo Brasil. Prefeituras, escritórios de advocacia e centros de assistência social são exemplos de entidades que poderão ser habilitadas para prestar parte dos serviços do INSS à sociedade. Como ponto positivo, essa medida proporciona um maior esvaziamento das outrora lotadas Agências da Previdência. No entanto, cabe destacar que o distanciamento criado entre a instituição e as pessoas tende a dificultar a execução de medidas preventivas e corretivas no sentido de combater as fraudes contra o dinheiro público.

Para que uma entidade possa firmar um Acordo de Cooperação Técnica junto ao INSS, a mesma deverá preencher uma série de requisitos necessários ao seu credenciamento como entidade parceira do Instituto. Esses requisitos compreendem uma série de documentos, como certidões negativas emitidas pela Receita Federal do Brasil e pela Procuradoria-Geral da Fazenda Nacional, comprovação de que não existe pendência da instituição com o INSS, registro no Cadastro Nacional de Pessoas Jurídicas (CNPJ), entre outros. Cabe destacar, ainda, que um ACT poderá ter uma vigência máxima de 60 [sessenta] meses [13].

Por fim, cabe destacar que a competência das entidades parceiras é restrita apenas ao protocolo dos requerimentos junto ao INSS, cabendo, ao Instituto, a atribuição exclusiva para a conclusão dos mesmos [13]. No entanto, a acordante deverá responder pela veracidade das informações prestadas, bem como dos documentos por ela recebidos, devendo ser responsabilizada, no caso de falhas ou erros que possam acarretar prejuízo ao INSS ou à população.

2.4 REDES NEURAIS ARTIFICIAIS – RNA

2.4.1 INTRODUÇÃO

Inspiradas no funcionamento da mente humana, as Redes Neurais Artificiais são modelos matemáticos projetados para simular a maneira como o cérebro realiza umas tarefas cotidianas. Para que atinjam um desempenho satisfatório, as Redes Neurais Artificiais utilizam uma maciça interconexão de “células” computacionais mais simples, os “neurônios artificiais”, também conhecidos como “unidades de processamento” [14].

2.4.2 HISTÓRICO

Os primeiros autores a realizar, em seus estudos, uma analogia entre o funcionamento das células nervosas e os processos eletrônicos foram o psiquiatra Warren McCulloch e o matemático Walter Pitts. Os trabalhos de McCulloch e Pitts resultaram na construção de uma unidade de processamento singular que ficou conhecida como o primeiro neurônio artificial [20]. O neurônio em questão era um modelo construído com resistores variáveis e amplificadores, que representavam as conexões sinápticas de um neurônio biológico.

A partir de então, vários pesquisadores passaram a dedicar seus estudos no desenvolvimento e aprimoramento de Redes Neurais. Cabe destacar os trabalhos de Frank Rosenblatt que, entre os anos de 1957 e 1958, construiu o Mark I – Perceptron, considerado o primeiro neuro-computador. O modelo idealizou o Perceptron, sendo capaz de reconhecer padrões simples. Outro trabalho de destaque foi desenvolvido por Widrow e Hoff, que consistiu numa regra de aprendizagem como extensão do Perceptron, conhecida como ADALINE.

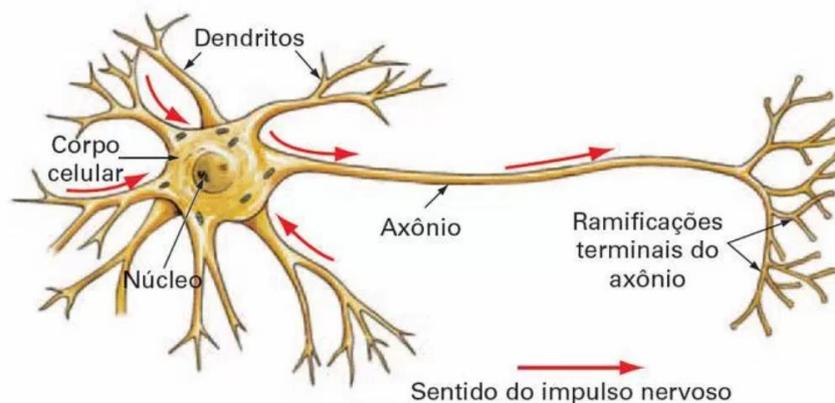
No entanto, por volta de 1969, pesquisadores constataram que o Perceptron só conseguia distinguir padrões linearmente separáveis. Essa constatação gerou grande desinteresse da comunidade acadêmica no estudo das RNAs. Apenas em 1982, com os trabalhos de John Hopfield e o desenvolvimento do algoritmo de treinamento *backpropagation*.

2.4.3 NEURÔNIO BIOLÓGICO

O neurônio biológico é considerado a unidade básica de transmissão e informações do sistema nervoso cerebral. O principal papel dos neurônios biológicos é a transmissão de informações, através da condução de impulsos elétricos. É formado, basicamente, por três partes: dendritos, corpo celular e axônio. Os dendritos são responsáveis por captar os impulsos advindos de outros neurônios, sejam estes pertencentes ao mesmo meio, ou a meios diferentes.

Já o corpo celular, tem como principal atribuição o processamento das informações captadas nos dendritos. É no corpo celular que as informações são coletadas e combinadas para que possam gerar novos impulsos elétricos que deverão ser transmitidos para outros neurônios. Por fim, temos o axônio, cuja missão é conduzir, através das ramificações denominadas conexões sinápticas, ou, simplesmente, sinapses, os impulsos elétricos para outros neurônios. A figura 4 traz uma representação gráfica de um neurônio biológico.

FIGURA 4 - NEURÔNIO BIOLÓGICO



FONTE: REPRODUZIDO DE <http://deeplearningbook.com.br/o-neuronio-biologico-e-matematico/>

2.4.4 NEURÔNIO ARTIFICIAL

Considerado a unidade básica de processamento das Redes Neurais Artificiais, o neurônio artificial consiste em uma representação matemática simplificada de um neurônio biológico. O funcionamento inspirado no cérebro humano, faz com que problemas complexos possam ser decompostos em tarefas mais simples, capazes de ser processada por esta singela estrutura computacional,

como coletar sinais em suas entradas, agrupá-los e transmiti-los na forma de resposta. O neurônio artificial é não-linear, o que torna seu uso bastante interessante para casos nos quais a entrada fornecida pelo meio físico tem um comportamento não-linear [14]. Ademais, os neurônios artificiais são capazes de emitir saídas contínuas.

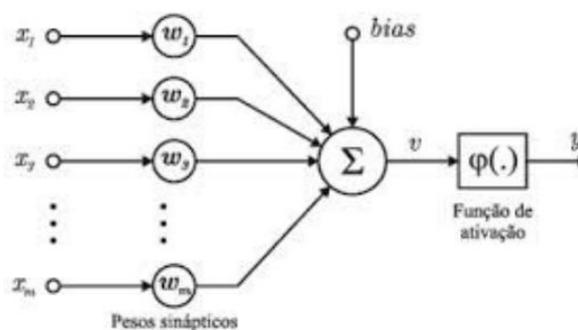
Na ilustração a seguir, podemos observar um modelo de neurônio artificial. São apresentadas entradas ($x_1, x_2, x_3, \dots, x_n$), representando os estímulos recebidos pelo neurônio, advindos do meio externo ou de outros neurônios. Também são ilustrados os pesos sinápticos ($w_1, w_2, w_3, \dots, w_n$), que significam o quão importante uma informação de entrada poderá ser, tendo um valor diretamente proporcional ao grau importância da informação.

Por fim, temos que o somatório do resultado obtido de todos os produtos das multiplicações efetuadas entre os valores de entrada e os pesos, subtraído o bias, gera um resultado v_i . Esse resultado será o parâmetro a ser utilizado pela função de ativação $\varphi(\cdot)$, a qual irá gerar a saída do neurônio.

$$v_i = \sum_{j=1}^n w_{ij} \cdot x_j - bias \quad [2.1]$$

Os pesos sinápticos conferem um grande potencial de adaptabilidade às Redes Neurais Artificiais, já que os mesmos são ajustados de acordo com as modificações do meio em que a Rede esteja envolvida [14]. A modificação dos pesos ocorre através de processos de treinamento da rede.

FIGURA 5 - NEURÔNIO ARTIFICIAL.



2.4.5 MULTI-LAYER PERCEPTRON – MLP

Multi-Layer Perceptron (MLP) é uma RNA formada por unidades sensoriais, também conhecidas como nós, aglutinadas em camadas, sendo uma camada de entrada, uma ou mais camadas intermediárias, denominadas camadas ocultas, e uma camada de saída. Em uma rede MLP, cada nó de uma camada deve estar conectado apenas aos nós da camada seguinte, não havendo, portanto, retroalimentação. Com isso, o sinal se propaga “para frente” através da rede, percorrendo camada por camada, até atingir a camada de saída.

As sinapses são representadas pelas conexões existentes entre os nós da rede. A camada de entrada é responsável por receber os impulsos que alimentarão o sistema. Assim, a quantidade de nós desta camada será estabelecida com base na dimensão das variáveis de entrada da RNA. A camada intermediária, por sua vez, confere a característica de não-linearidade ao sistema, tornando as redes MLP capazes de lidar com problemas do mundo real. A resposta da rede, que representa o valor desejado, será obtida na camada de saída.

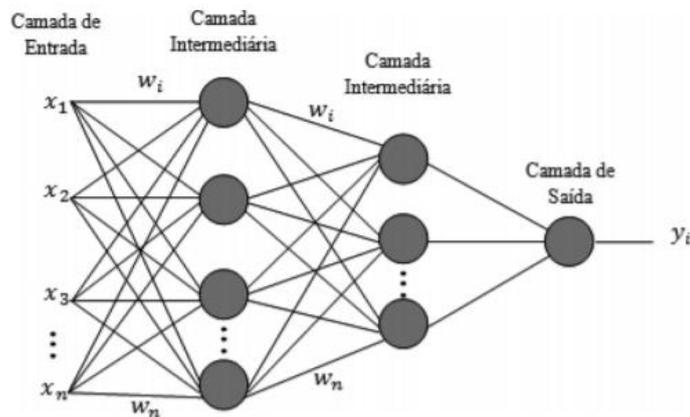
Amplas são as possibilidades de aplicação da arquitetura MLP. Sua alta adaptabilidade torna possível sua utilização face a vários tipos de problemas do mundo real, nos mais variados campos de estudo das ciências. A Figura 6 traz uma representação de uma rede MLP. São exemplos de problemas que podem ser solucionados com o uso de MLP os seguintes:

- Séries Temporais.
- Reconhecimento de Padrões.
- Robótica.
- Processamento de Imagens e de Sinais.
- Análise de Mercado Financeiro.
- Visão Computacional.

O treinamento das redes MLP é feito de forma supervisionada. Assim, faz-se necessária a utilização de um algoritmo capaz de construir um conjunto de pesos

ótimos a serem aplicados nas conexões da rede. No entanto, a existência de camadas intermediárias escondidas acrescenta um grau de complexidade ao treinamento, já que, por serem camadas ocultas, não se tem conhecimento sobre o erro a ser utilizado no ajuste dos pesos. Devido a esse comportamento, foi desenvolvido um algoritmo que propaga, de forma recursiva, o sinal erro encontrado na saída da rede no sentido contrário ao das conexões sinápticas. Tal algoritmo é conhecido como *backpropagation*, ou algoritmo de retropropagação de erro.

FIGURA 6 - EXEMPLO DE UMA ARQUITETURA MLP.



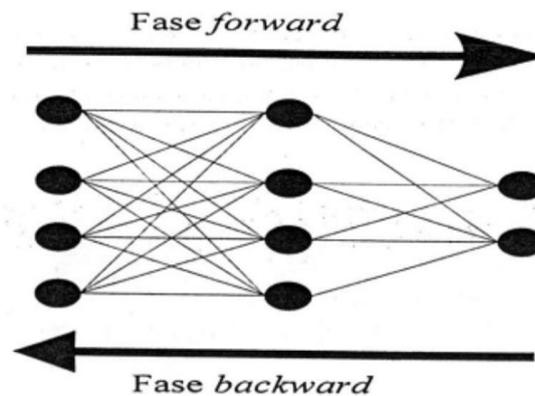
FONTE: REPRODUZIDO DE http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-77862016000100024

O algoritmo de *backpropagation* funciona, basicamente, em duas etapas: na primeira delas, os sinais de entrada são propagados no sentido das conexões sinápticas, percorrendo camada por camada até a saída do sistema. Ao final do percurso, calcula-se o erro, sem que seja realizado nenhum ajuste nos pesos. Na segunda etapa, o erro obtido é propagado da camada de saída para a camada de entrada, de forma recursiva. Nessa etapa, os pesos sinápticos são ajustados para que a resposta da rede se aproxime da condição desejada. A esse processo de aprendizagem, dá-se o nome de aprendizagem por retropropagação.

2.4.5.1 ALGORITMO BACKPROPAGATION

O aprendizado por retropropagação de erro consiste na correção dos pesos sinápticos através da aplicação recursiva do erro obtido na saída da RNA, percorrendo a rede no sentido reverso ao das conexões sinápticas. O processo é realizado em duas fases, conhecidas como fase forward (“para frente”) e fase backward (“para trás”). A fase forward é executada para que se obtenha a saída da rede. Já a fase backward consiste na atualização dos pesos das sinapses. A figura 7 traz o sentido em que as duas fases percorrem a rede.

FIGURA 7 - FASES DO ALGORITMO BACKPROPAGATION.



FONTE: REPRODUZIDO DE [21].

Fase forward

A Fase *forward* consiste na propagação do sinal de entrada por todas as camadas da RNA. O sinal percorre a rede no sentido das conexões sinápticas, até a saída da mesma. É nessa fase que se obtém o erro do valor de saída, que será utilizado como parâmetro de ajuste dos pesos a ser realizado na fase backward. A equação 2.2, representa matematicamente como é obtido o valor do sinal de cada neurônio na camada de saída, a saber:

$$y_i^2 = f(\text{net}_i^2) \quad [2.2]$$

Da equação acima, temos que:

- y_i^2 - Valor do sinal do neurônio i da camada 2, saída para uma RNA com uma camada escondida;
- f - Função de ativação da camada 2;
- net_i^2 - Sinal de entrada do neurônio i da camada 2.

A seguir, a equação 2.3 demonstrará o cálculo de erro para cada neurônio na saída:

$$e_i(n) = (d_i - y_i^2) \quad [2.3]$$

Sendo $e_i[n]$ o valor do erro do neurônio i pertencente à camada de saída (n). O valor desejado para o neurônio i é d_i e y_i^2 é o valor de saída do neurônio 2, obtido através da equação 2.2.

Fase *backward*

A segunda fase do processo de aprendizagem por retropropagação, consiste na aplicação recursiva do erro obtido na saída, que percorre a rede no sentido contrário ao das conexões sinápticas. Aqui, os pesos devem ser ajustados, de acordo com a equação 2.4

$$w_{ij}^m(novo) = w_{ij}^m(antigo) + \alpha \delta_i^m f^{m-1}(net_j^{m-1}) \quad [2.4]$$

Onde w_{ij}^m é o peso das sinapses entre os neurônios i , que emite o sinal, e o neurônio j que recebe o sinal. α representa a taxa de aprendizagem e δ_i^m determina a sensibilidade do neurônio emissor do sinal. Tal sensibilidade é obtida através de um cálculo que leva em consideração duas situações. Na primeira delas, o neurônio j receptor do sinal, já se encontra na camada de saída da rede. Nesse caso, é utilizada a equação 2.5, descrita a seguir:

$$\delta_i^2 = f^{2'}(net_i^2) \cdot e_i(n) \quad [2.5]$$

Já na segunda hipótese, o neurônio j não se encontra na camada de saída da rede. Assim, a sensibilidade do neurônio será calculada de forma recursiva, da última para a primeira camada, com base na equação 2.6 abaixo:

$$\delta_j^{m-1} = f^{m-1'}(net_j^{m-1}) \sum_{i=1}^N w_{ij}^m \delta_i^m \quad [2.6]$$

Nesta equação, $f^{m-1'}(net_j^{m-1})$ significa a derivada da função de ativação da camada emissora do estímulo, o índice i do somatório em questão, representa a quantidade de neurônios da camada receptora do estímulo, a qual encontra-se imediatamente a frente e possui N neurônios com sensibilidade δ_i^m .

3 METODOLOGIA

O capítulo em tela traz as etapas percorridas para que os objetivos listados no capítulo inicial deste trabalho fossem atingidos. Ademais, também serão apresentadas as aplicações dos conhecimentos teóricos explanados no capítulo 2 na execução dos estudos propostos nesta monografia.

3.1 BASE DE DADOS

Para a construção do modelo, foram selecionadas variáveis de entrada que representam características recorrentes nos processos fraudulentos detectados no INSS. A escolha destas variáveis tomou como base as várias operações realizadas pela Polícia Federal, em parceria com o INSS, para a desarticulação de quadrilhas que atuam por todo o Brasil. Assim, para que as variáveis de entrada se possam retratar os diversos *modus operandi* dos criminosos, as amostras foram divididas em dois grupos, quais sejam os benefícios previdenciários e os benefícios assistenciais.

Tal divisão se deve ao fato de que existem características inerentes aos requerimentos de benefícios pertencentes a um mesmo grupo. Com isso, a análise de risco fica mais focada nos pontos sensíveis de cada requerimento, de acordo com o grupo a que este requerimento pertença.

3.1.1 GRUPO 1- BENEFÍCIOS ASSISTENCIAIS

Nos requerimentos pertencentes ao grupo dos benefícios assistenciais, foram selecionadas quatro informações cujo comportamento pode indicar fortemente a possibilidade de ocorrência de fraude contra a Previdência. As variáveis são:

- Grupo Familiar: Quantidade de pessoas que residem com o requerente. Em grande parte dos casos, os fraudadores omitem informações dos familiares que residem com o requerente, para que não sejam detectados rendimentos financeiros acima do permitido pela legislação.
- Renda *per capita*: De acordo com o parágrafo 3º do Artigo 20 da Lei Orgânica da Assistência Social (LOAS), caso a família da pessoa idosa ou com

deficiência possua renda inferior a $\frac{1}{4}$ do salário-mínimo, a mesma poderá ter direito a um benefício assistencial. Assim, caso seja detectada uma renda superior ao limite legal, existe a possibilidade de que este requerimento configure uma ilegalidade. Logo, o prévio conhecimento da condição financeira do requerente constitui informação fundamental para a identificação de possíveis fraudes contra a Previdência Social.

- Unidade da Federação (UF) do requerimento: Não obstante a atuação de quadrilhas especializadas em crimes previdenciários ocorrer em todo o território nacional, alguns estados possuem, historicamente, uma maior incidência de casos. Estados como Maranhão e Rio de Janeiro, por exemplo, são amplamente conhecidos no ambiente corporativo do INSS, bem como em relatórios emitidos por instituições de controle, como possuindo maior ocorrência de violações legais.
- Acordo de Cooperação Técnica: Por se tratar de inovação criada pelo INSS Digital, onde os requerimentos poderão ser protocolados fora dos domínios do INSS, sem a presença de servidores públicos do órgão, os ACTs podem representar um risco à instituição, já que, apesar de conferir maior flexibilidade e capilaridade, promove a realização de pedidos sem que haja o crivo dos servidores públicos da instituição. Essa flexibilização poderá encorajar os fraudadores, já que facilita a cooptação de pessoas visando burlar as leis previdenciárias.

As duas primeiras variáveis foram consideradas com base nas observação de Notas Técnicas emitidas pelo Ministério do Desenvolvimento Social e Agrário (MDSA), em especial a Nota Técnica nº 7/2017, emitida pelo Gabinete da Secretaria de Avaliação e Gestão da Informação do citado Ministério. Neste documento, foram consolidadas as conclusões da análise de uma massa de dados de cerca de 4.561.620 benefícios assistenciais, visando a detecção de irregularidades na manutenção destes benefícios. Neste relatório, os técnicos do MDSA relatam que, entre os parâmetros utilizados para a detecção de irregularidades, estão o Grupo Familiar e as respectivas informações de renda *per capita* deste.

A terceira variável foi escolhida com base em relatórios e matérias jornalísticas que apontam para uma maior atuação de quadrilhas em determinados

estados do Brasil. De posse dessa informação, os servidores públicos normalmente tendem a ter mais cuidado na análise de requerimentos advindos de regiões com alta incidência. Segundo dados da Delegacia de Repressão a Crimes Previdenciários do Maranhão, entre 2014 e 2016, foram executadas mais de 18 operações policiais contra ações de má-fé. Ainda, de acordo com o Ministério Público Federal [MPF], em 2016, mais de 1500 inquéritos foram abertos para verificar crimes contra o INSS [15].

Por fim, a quarta variável foi selecionada por sua própria concepção. Os ACTs preveem um distanciamento maior entre a população e a previdência. Assim a RNA a ser construída nesse projeto representa uma ferramenta de controle que, usando os artifícios da Inteligência Artificial, poderá fiscalizar esse novo fluxo criado pelo INSS Digital.

3.1.2 GRUPO 2- BENEFÍCIOS PREVIDENCIARIOS

No caso dos requerimentos pertencentes ao grupo dos benefícios previdenciários, foram selecionadas as seguintes variáveis, que, assim como nos benefícios assistenciais, representam informações relevantes ao ser avaliado o risco de fraudulência de um requerimento. As variáveis estão listadas a seguir:

- **Espécie:** Através da espécie, podemos identificar qual o tipo de benefício está sendo requerido. Benefícios como salário-maternidade e auxílio reclusão, historicamente, possuem grande incidência de fraudes. Assim, conhecer o tipo de benefício que está sendo requerido pode alertar o servidor sobre a possibilidade de ocorrência de crimes.
- **Unidade da Federação (UF) do requerimento:** De forma semelhante ao grupo apresentado na seção 3.1.1, conhecer a procedência geográfica do requerimento pode fazer com que o servidor público eleve o seu nível de atenção ao analisar um pedido de benefícios.
- **Acordo de Cooperação Técnica:** Também segue a mesma idéia da seção 3.1.1. O requerimento proveniente de ACT pode despertar um grau de atenção maior quando da análise da demanda.

- Natureza: Aqui classificaremos os requerimentos de acordo com a natureza da atividade laboral do solicitante, que pode ser urbana ou rural. Essa divisão se dá porque, em algumas regiões do país, há uma maior incidência de fraudes em benefícios rurais, em outros locais, por sua vez, a quantidade de benefícios urbanos fraudulentos é maior. Assim, esse parâmetro, ao ser combinado com outros parâmetros de entrada, poderá ocasionar em um resultado que represente um maior nível de alerta com relação à possibilidade de ocorrência de irregularidades.

As duas primeiras variáveis são escolhidas com base em relatórios de órgãos de controle e matérias jornalísticas em que é possível detectar que, dependendo da espécie do benefício ou do lugar de procedência do requerimento será maior a probabilidade de fraude, haja vista dados citados na seção 3.1.1, onde o estado do Maranhão contém o maior número de ocorrências de crimes contra a Previdência Social. Ademais, os requerimentos de salário-maternidade, por exemplo, contém número expressivo de irregularidades, como podemos observar em [15], [16] e [17]

A terceira variável foi selecionada pelos mesmos motivos já explanados na seção 3.1.1. A quarta variável também representa um dado que inspira cuidados do técnico responsável pela análise da demanda. Em alguns casos, como em [5] e [6], sindicatos rurais de uma determinada localidade estão funcionando como verdadeiras fábricas de requerimentos fraudulentos. Em outras situações, como em [16] e [17], são os intermediários dos grandes centros urbanos que cooptam pessoas para participar do esquema criminoso. Assim, a natureza do benefício, combinada a outros fatores, poderá aumentar o nível de risco das demandas previdenciárias.

3.2 PROCESSAMENTO DE DADOS

As amostras de valores para as variáveis listadas nas seções 3.1.1 e 3.1.2 serão obtidas em um conjunto de 40 requerimentos, sendo 20 do grupo 1 e 20 do grupo 2. Os dados serão advindos de processos mantidos no acervo da própria

instituição e, por questões de confidencialidade, não serão revelados os dados pessoais dos solicitantes.

3.2.1 CONVERSÃO DE DADOS

Para que os dados possam ser submetidos às RNA, é preciso que os mesmo sejam descritos em uma notação capaz de ser utilizada como estímulos de entrada da rede. Com isso, foram definidos valores numéricos e booleanos para as variáveis que representam valores lógicos, ou para aquelas que representaram faixas de valores. As tabelas a seguir apresentam os possíveis valores das variáveis e seu correspondente número ou valor lógico. Cabe destacar que a tabela 2 divide a renda per capita em faixas de valores, com base na divisão feita na Nota Técnica nº 7 do MDSA [18]. A tabela 4, por sua vez, associa os tipos de benefícios com os seus códigos respectivos, com base na tabela de códigos dos benefícios do INSS [19].

TABELA 1 - CONVERSÃO DE DADOS DA VARIÁVEL GRUPO FAMILIAR.

Grupo Familiar	
Valor real	Valor numérico correspondente
Apenas o requerente	1
De 2 a 4 pessoas [considerando o requerente]	2
Acima de 4 pessoas [considerando o requerente]	3

FONTE: ELABORADA PELO AUTOR.

TABELA 2 - CONVERSÃO DE DADOS DA VARIÁVEL RENDA *PER CAPITA*.

Renda per capita	
Valor real	Valor numérico correspondente
Até $\frac{1}{4}$ do salário mínimo	1
De $\frac{1}{4}$ a $\frac{1}{2}$ salário mínimo	2
Acima de $\frac{1}{2}$ salário mínimo	3

FONTE: ELABORADA PELO AUTOR.

TABELA 3 - CONVERSÃO DE DADOS DA VARIÁVEL ACT.

Acordo de Cooperação Técnica [ACT]	
Valor real	Valor booleano correspondente
Solicitação não proveniente de ACT	FALSE
Solicitação proveniente de ACT	TRUE

FONTE: ELABORADA PELO AUTOR.

A tabela 4 apresenta a conversão dos valores da variável Espécie para equivalentes numéricos. Esses códigos numéricos são utilizados cotidianamente na análise dos requerimentos apresentados ao INSS. A associação entre cada tipo de benefício previdenciário e um dos códigos presentes na citada tabela encontra-se disponível em [19]. Algumas espécies não são mais ofertadas à população, como é o caso do benefício 07 – Aposentadoria por Idade do Trabalhador Rural e por isso esses tipos foram desconsiderados neste trabalho. Os benefícios assistenciais, previstos na seção 2.2 também possuem códigos numéricos correspondentes. No entanto, durante a seleção das características mais relevantes na determinação do risco de fraude de um requerimento apresentado junto ao INSS, foi verificado que, nos casos de benefícios assistenciais, a espécie não é um fator determinante para a

atribuição de um indicador de risco, já que os benefícios possuem regras de aquisição semelhantes, sendo, inclusive, previstos na mesma lei, a saber a Lei 8.742 de 7 de dezembro de 1993, conhecida como a Lei Orgânica da Assistência Social (LOAS).

TABELA 4 - CONVERSÃO DE DADOS DA VARIÁVEL ESPÉCIE.

Espécie de Benefícios Previdenciários	
Valor real	Valor Numérico Correspondente
Aposentadoria por Idade	41
Aposentadoria por Tempo de Contribuição	42
Aposentadoria por Invalidez	32
Aposentadoria Especial	46
Auxílio-Doença	31
Salário-Maternidade	80
Auxílio-reclusão	25
Pensão por Morte	21
Auxílio-Acidente	94

FONTE: EXTRAÍDO DE <http://www.normaslegais.com.br/guia/clientes/codigo-beneficios-previdencia-social.htm>

As demais variáveis a serem consideradas serão levadas pelos seus valores literais, já que o *Weka* admite que as variáveis possam ser inseridas textualmente, em formato de *string*.

3.3 ARQUITETURA MLP NA IDENTIFICAÇÃO DE FRAUDES EM REQUERIMENTOS DE BENEFÍCIOS PREVIDENCIÁRIOS

A configuração de uma rede MLP a ser utilizada nos experimentos propostos nesta monografia foi feita através do *software Weka*, versão 3.8.3. A seleção da referida arquitetura é feita através do módulo *Explorer*, na aba *Classify*. Em seguida, foram desenvolvidos, utilizando a ferramenta *Notepad++*, versão 7.5.9, dois arquivos contendo as amostras de informação a serem utilizadas no processo de treinamento e validação da rede. Esses arquivos devem ser salvos com a extensão *.arff*. A extensão se refere à expressão *Attribute-Relation File Format* (ARFF), que representa o formato dos arquivos a serem abertos pelo *Weka*. Como numa linguagem de programação, o arquivo *.arff* deve seguir regras específicas de sintaxe e semântica para que o software simulador possa aceitá-lo como entrada.

O arquivo pode ser criado também através de uma planilha *Excel*. Para tanto, devemos salvar o arquivo como com a extensão *.csv*, para que, em seguida, o mesmo possa ser convertido para *.arff*, utilizando alguns *softwares* como o *ExceltoARFF*. A figura 7 apresenta o formato do arquivo *.arff* criado para ser submetido ao *Weka*. Insta salientar que a criação de dois arquivos se fez necessária devido à divisão das amostras entre benefícios assistenciais e benefícios previdenciários.

Após a seleção da arquitetura a ser utilizada e a criação do arquivo com os dados, foi preciso ajustar alguns parâmetros da rede para que a mesma procedesse à execução dos testes de forma esperada. Os parâmetros configurados são os seguintes

- *learningRate*: Representa a taxa de aprendizagem, a qual foi atribuído o valor 0,3. Cabe destacar que este foi o valor padrão escolhido pelo *Weka*.
- *seed*: Configurada inicialmente com o valor 1, esse parâmetro precisa ser modificado antes de realizar as simulações. Se o valor da *seed* se mantiver constante, as simulações irão apresentar os mesmos resultados.
- *Percentage split*: Representa o percentual da amostra que será utilizado para o treinamento da rede. O valor foi ajustado para 50%. Assim 50% dos valores serão utilizados para treinamento da rede, e 50% serão utilizados para a validação da rede em questão.

FIGURA 8 - ARQUIVO DE TEXTO COM OS DADOS DE BENEFÍCIOS ASSISTENCIAIS A SEREM INSERIDOS NO WEKA.

```

1 @relation assistenciais
2
3 @attribute grupoFamiliar real
4 @attribute renda real
5 @attribute ufRequerimento {Acre, Alagoas, Amazonas, Amapa, Bahia, Ceara, DistritoFederal, EspiritoSanto, Goias, Maranhao,
6 MatoGrosso, MatoGrossodoSul, MinasGerais, Para, Paraiba, Parana, Pernambuco, Piaui, RiodeJaneiro, RioGrandedoNorte, RioGrandedoSul,
7 Rondonia, Roraima, SantaCatarina, SaoPaulo, Sergipe, Tocantins}
8 @attribute act {TRUE, FALSE}
9 @attribute fraude {yes, no}
10
11 @data
12 1,1,Maranhao,FALSE,yes
13 1,3,Maranhao,FALSE,yes
14 2,1,RiodeJaneiro,FALSE,no
15 1,3,Maranhao,FALSE,yes
16 2,2,RioGrandedoSul,TRUE,no
17 1,1,Maranhao,TRUE,yes
18 2,3,Pernambuco,TRUE,yes
19 2,1,Maranhao,TRUE,yes
20 3,2,Parana,TRUE,no
21 1,1,RiodeJaneiro,FALSE,yes
22 1,3,RiodeJaneiro,FALSE,yes
23 1,1,EspiritoSanto,FALSE,yes
24 3,1,EspiritoSanto,FALSE,yes
25 3,1,Maranhao,TRUE,yes
26 3,1,Bahia,FALSE,no
27 2,1,Bahia,FALSE,no
28 3,3,Ceara,FALSE,yes
29 1,1,Bahia,TRUE,no
30 1,1,RioGrandedoSul,TRUE,no
31 1,1,Parana,FALSE,no
32 1,1,MinasGerais,TRUE,yes
33 1,1,Tocantins,FALSE,yes
34 1,3,Tocantins,TRUE,yes
35 1,3,Tocantins,FALSE,yes
36 1,1,Maranhao,TRUE,yes
37 1,3,Ceara,TRUE,yes
38 2,1,Pernambuco,TRUE,no
39 2,1,SaoPaulo,FALSE,no
40 2,1,SaoPaulo,FALSE,no
41 1,1,Parana,FALSE,no

```

FONTE: ELABORADO PELO AUTOR.

Uma vez configurados os parâmetros da RNA, foram submetidos como amostra, dois arquivos *.arff*, contendo os dados de benefícios previdenciários e assistenciais colhidos no acervo do INSS. Esses arquivos foram utilizados para o treinamento da rede, bem como validação dos resultados, como descrito no parâmetro Percentage split acima definido.

Para fins de avaliação do comportamento da RNA frente a estas duas amostras, observaremos as informações listadas no *Classifier Output*, localizado na aba *Classify* do *Weka*. Esse console apresenta uma descrição do processo de aprendizagem e validação da rede, listando dados importantes para a verificação da consistência da RNA após a realização do experimento. Para tanto, serão anotados em destaque os valores dos seguintes parâmetros:

- *Mean absolute error* - EMA.
- *Root mean squared error* - REMQ.

Também é de relevante importância observar os dados referentes ao processo de classificação da RNA após o seu treinamento, haja vista a necessidade

de verificarmos se as redes MLP podem ser utilizadas para o problema apresentado neste trabalho. Com isso, dois valores passaram a ser observados, sob o seguinte parâmetro:

- *Correctly Classified Instances.*
- *Incorrectly Classified Instances.*

Estes parâmetros representam numericamente, a quantidade de casos em que a RNA classificou os dados corretamente ou incorretamente. Com eles, podemos constatar se a forma de aprendizagem da arquitetura MLP representa uma alternativa plausível para o fim proposto nesta monografia. Ademais, é possível, com o conhecimento dos valores destas variáveis e em associação a outros dados, quantificar os casos incorretos e atuar de forma a minimizar a sua ocorrência.

4 RESULTADOS

A análise dos resultados apresentados pela arquitetura MLP quando aplicada para solucionar o problema proposto neste trabalho é de grande relevância para que possamos atestar se a mesma pode ser utilizada como alternativa na constatação de fraudes em requerimentos de benefícios previdenciários e assistenciais protocolados junto ao INSS. Para tanto, a RNA precisa ser capaz de, dado um conjunto de informações baseadas em características singulares de cada requerimento, identificar se há indícios de irregularidades.

4.1 ANÁLISE DOS REQUERIMENTOS DE BENEFÍCIOS ASSISTENCIAIS

Para esta análise, foi apresentado um conjunto de dados de 50 requerimentos de benefícios assistenciais presentes no acervo do INSS, para que a RNA apresentasse quais podem ser fraudulentos e quais não contém risco de irregularidades. Como já descrito na seção 3.3, ajustamos a variável Percentage Split para 50. Assim, 50% dos dados da amostra foram utilizados para o treinamento da rede, enquanto que os 50% restantes foram utilizados para a validação da mesma. Impende salientar, ainda, que os dados utilizados para treinamento e validação do modelo foram escolhidos aleatoriamente, dentre os valores contidos no espaço amostral submetido à RNA.

Após a execução de 50 simulações através do *Menu Experimenter*, do *Weka*, alguns dados relevantes puderam ser conhecidos, através da aba *Analyse*, opção *Experimenter*. O teste de performance realizado revelaram um percentual de acerto de 75% dos casos apresentados à rede, o que revela que, das 30 amostras apresentadas para fins de validação da arquitetura, cerca de 24 foram classificadas corretamente. A margem de erro dessa estimativa, representada pelo desvio padrão, é da ordem de 6,40%, significando, assim, que, cerca de 6 amostras podem ter sido classificadas erroneamente. Cabe destacar, ainda, que o nível de significância, que representa o grau de confiança dos dados apresentados, é da ordem de 10%. Outros valores relevantes a serem destacados são os seguintes:

- *Mean absolute error*: O erro médio absoluto traduz a possível existência de um comportamento tendencioso do modelo em análise. Para que a RNA possa ser considerada isenta, o EMA deve ser um valor próximo de 0 [zero]. Na rede em questão, o valor obtido após as 50 simulações foi 0,184.
- *Root mean squared error [REMQ]*: O erro médio quadrático funciona como uma medida de qualidade das previsões realizadas por um determinado modelo de estimativas. No caso em tela, o REMQ poderá ser considerado para a escolha de uma determinada arquitetura de RNA para a solução do problema proposto. Para a arquitetura MLP, o REMQ foi 0,434. Aqui, quanto mais próximo de 0 for o valor, melhor será a previsibilidade do modelo.
- *Correctly Classified Instances*: Quantidade de amostras classificadas corretamente. Após as simulações, foram verificados cerca de 24 casos classificados corretamente, em um universo de 30 amostras.
- *Incorrectly Classified Instances*: Quantidade de amostras classificadas incorretamente, após as simulações, foi verificado que cerca de 6 amostras foram classificadas incorretamente.

Impende destacar, ainda, que para o modelo escolhido, não foram detectados casos que ficaram sem classificação. O parâmetro *number_unclassified* possui valor nulo, significando para cada valor de entrada apresentado na fase de validação da RNA, o modelo mostrou-se capaz de prever um comportamento, com base nos dados de entrada.

4.2 ANÁLISE DOS REQUERIMENTOS DE BENEFÍCIOS PREVIDENCIÁRIOS

De forma análoga ao disposto na seção 4.1, foram executadas 50 simulações, com base em requerimentos de benefícios previdenciários, extraídos do acervo do INSS. Alguns destes requerimentos possuem indícios de irregularidade, enquanto que outros são solicitações onde não foram verificados traços característicos de uma demanda fraudulenta. Aqui, a variável *Percentage split* foi ajustada para 50, ou seja, 50% dos dados foram utilizados para treinamento, ao

passo que os 50% restantes foram utilizados para a validação do modelo proposto. Cabe destacar, ainda, que tanto o conjunto de dados para treinamento, quanto o conjunto de dados de validação foram escolhidos aleatoriamente, dentre as amostras apresentadas ao sistema.

O arquivo *beneficiosPrevidenciarios.arff* foi criado para alimentar o *Weka* com os dados obtidos do acervo da instituição. Assim, como na seção 4.1, alguns parâmetros obtidos após a realização das simulações merecem destaque. As simulações foram realizadas no módulo *Explorer* do *Weka*. O teste de performance constatou um percentual de acerto da ordem de 86%, com desvio padrão da ordem de 6.5% e nível de significância de 10%. Isso nos dá uma média de, aproximadamente, 26 acertos em um grupo de 30 amostras. A seguir, são descritos alguns parâmetros de grande relevância para a análise das simulações:

- *Mean absolute error*: O erro médio absoluto obtido foi de 0.16%.
- *Root mean squared error*: REMQ com valor de 0.32%
- *Correctly Classified Instances*: Uma média de 26 acertos em 30 amostras.
- *Incorrectly Classified Instances*: Com valor médio de 4 erros em 30 amostras.

Cabe salientar, por fim, que para a arquitetura escolhida, não foram detectados casos que ficaram sem classificação, haja vista o valor da variável *number_unclassified* ser nulo. Tal comportamento denota que, para cada valor de entrada da RNA, haverá uma classificação final correspondente aos estímulos informados.

5 CONSIDERAÇÕES FINAIS

5.1 CONCLUSÕES

Este trabalho teve como norte principal propor o uso dos conceitos de Inteligência Artificial no combate às fraudes e erros nas análises das demandas apresentadas pela população brasileira ao Instituto Nacional do Seguro Social. A idéia central desta monografia busca inspiração nas inovações implantadas através da execução do projeto “INSS Digital: Uma nova forma de atender”, que consistem em uma série de mudanças no fluxo de atendimento ao público do INSS, visando uma maior otimização e economia em seus processos.

Por se tratar de uma idéia gestada ainda no decorrer deste trabalho, foi utilizada, como modelo de RNA, a arquitetura *Multilayer Perceptron*. A escolha dessa arquitetura se deu em virtude de seu alto potencial computacional, bem como de sua consolidação na literatura específica como um dos modelos de redes neurais mais utilizados atualmente. Nos experimentos realizados, foram introduzidas na rede dados referentes a características marcantes existentes nos requerimentos de benefícios junto à Previdência Social brasileira. Através dos valores das características apresentadas, a RNA pode ser capaz de identificar se há ou não a possibilidade de que um dado pedido seja eivado de vícios.

Para tanto, os requerimentos foram divididos em dois grandes grupos, a saber: benefícios assistenciais e benefícios previdenciários, os quais se encontram descritos nas seções 3.1.1 e 3.1.2, respectivamente. Ademais, as simulações foram executadas através do *Weka*, versão 3.8.3, utilizando os seus módulos para treinamento e validação da rede MLP em questão.

Após as simulações, foram destacados alguns parâmetros relevantes para que possa ser possível atestar se a proposta de uso de uma rede MLP se mostra plausível para o problema proposto. Esses parâmetros estatísticos podem apontar para um comportamento aceitável ou não da RNA, quando a mesma se depara com o problema citado.

5.2 TRABALHOS FUTUROS

A partir da proposta sugerida neste trabalho, vários são os projetos que podem vir a surgir, buscando otimizar a idéia gestada nesta monografia, bem como buscando aplicar os conceitos de RNA em outras soluções no âmbito do INSS. A implantação de novos fluxos de atendimento no INSS traz consigo uma série de desafios que podem vir a ser solucionados através do uso da Inteligência artificial [IA]. Mais especificamente, podem ser compreendidos como linhas de pesquisa a serem exploradas em trabalhos futuros as seguintes:

- Integração das informações obtidas através das simulações realizadas neste trabalho, com os sistemas corporativos do INSS, visando colocar em prática a solução prevista nesta monografia.
- Normalização dos dados a serem inseridos na arquitetura proposta, visando uma maior confiabilidade da solução ora apresentada.
- Comparação da arquitetura MLP com outros modelos de redes neurais, visando encontrar uma arquitetura que tenha comportamento mais robusto frente ao problema proposto nesta monografia.
- Aplicação de Inteligência Artificial na análise e concessão de benefícios previdenciários, de forma a automatizar todo o fluxo de reconhecimento de direitos, reduzindo a manualização que gera grande morosidade para a conclusão dos requerimentos.
- Desenvolvimento de uma ferramenta que integre, com inteligência computacional, os dados dos cidadãos, constantes nos diversos órgãos de controle da república, visando a criação de um banco de dados unificado, capaz de verificar se uma determinada pessoa já possui, por exemplo, histórico de processos judiciais relacionados ao cometimento de crimes contra a Previdência Social.

REFERÊNCIAS

- (1) SECRETARIA DA PREVIDÊNCIA. RGPS: Previdência Social fecha 2017 com déficit de R\$ 182,4 bilhões. Brasília, 2018. Disponível em: <http://www.previdencia.gov.br/2018/01/rgps-previdencia-social-fecha-2017-com-deficit-de-r-1824-bilhoes/>. Acesso em: 01 set. 2018.
- (2) VALENTE, Gabriela. [O Globo Economia]. Previdência gasta R\$ 56 bilhões por ano com fraudes e erros, estima o TCU. Brasília, 2017. Disponível em: <https://oglobo.globo.com/economia/previdencia-gasta-56-bilhoes-por-ano-com-fraudes-erros-estima-tcu-21725551> . Acesso em: 01 set. 2018.
- (3) Portal INSS. INSS Digital: Uma nova forma de atender ao cidadão. Brasília 2017. Disponível em <https://www.inss.gov.br/inss-digital-nova-forma-de-atender-aos-segurados/> Acesso em: 02 set. 2018.
- (4) Portal INSS. Rede de Atendimento. Brasília, 2018. Disponível em: <https://www.inss.gov.br/servicos-do-inss/rede-de-atendimento/> . Acesso em: 02 set. 2018.
- (5) ESTARQUE, Thays. [G1 Pernambuco]. PF prende vereador e gerente de agência do INSS suspeitos de fraude. Recife, 2016. Disponível em: <http://g1.globo.com/pernambuco/noticia/2016/09/pf-prende-vereador-e-gerente-de-agencia-do-inss-suspeitos-de-fraude.html> . Acesso em: 01 set. 2018.
- (6) AGUIAR, Priscilla. [Portal FolhaPE]. Suspeitos de fraude milionária em aposentadorias rurais são presos pela PF em Pernambuco. Recife, 2017. Disponível em <https://www.folhape.com.br/noticias/noticias/cotidiano/2017/08/09/NWS,37362,70,449,NOTICIAS,2190-SUSPEITOS-FRAUDE-MILIONARIA-APOSENTADORIAS-RURAISSAO-PRESOS-PELA-PERNAMBUCO.aspx> . Acesso em 03 set. 2018.
- (7) SECRETARIA DA PREVIDÊNCIA. COMBATE ÀS FRAUDES: Força-Tarefa desarticula quadrilha em Pernambuco. Brasília, 2016. Disponível em: <http://www.previdencia.gov.br/2016/05/combate-a-fraudes-forca-tarefa-desarticula-quadrilha-em-pernambuco/> Acesso em: 05 set. 2018.
- (8) SECRETARIA DA PREVIDÊNCIA. COMBATE ÀS FRAUDES: Operação em São Paulo prende 12 pessoas por fraude previdenciária. Brasília, 2018. Disponível em: <http://www.previdencia.gov.br/2018/04/combate-as-fraudes-operacao-em-sao-paulo-prende-12-pessoas-por-fraude-previdenciaria/>. Acesso em: 06 set. 2018.
- (9) BRASIL. Lei nº 8.213 de 24 de julho de 1991. Dispõe sobre os Planos de Benefícios da Previdência Social e dá outras providências. Diário Oficial da União. Brasília 25 de julho de 1991. Disponível em: http://www.planalto.gov.br/ccivil_03/Leis/L8213compilado.htm. Acesso em: 10 set. 2018.

- (10) BRASIL. Constituição da República Federativa do Brasil de 1988. Diário Oficial da União. Brasília, 5 de Outubro de 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicaocompilado.htm . Acesso em: 12 de set. de 2018.
- (11) BRASIL. Lei nº 8.742 de 7 de dezembro de 1993. Dispõe sobre a organização da Assistência Social e dá outras providências. Diário Oficial da União. Brasília, 8 de dezembro de 1993. Disponível em: http://www.planalto.gov.br/ccivil_03/LEIS/L8742compilado.htm . Acesso em: 12 de set. de 2018.
- (12) MARCHESINI, Lucas; SIMÃO, Edna. [Valor Econômico]. Benefícios pagos pelo INSS representam mais de 25% do PIB em 500 municípios. Brasília, 2017. Disponível em: <https://www.valor.com.br/brasil/4832362/beneficios-pagos-pelo-inss-representam-mais-de-25-do-pib-em-500-municipios> . Acesso em: 15 de set. de 2018.
- (13) BRASIL. Anexo ao Memorando-Circular Conjunto nº 20 /DIRBEN/DIRAT/INSS de 4 de Julho de 2017. Guia prático para a celebração de Acordo de Cooperação Técnica – ACT. Brasília, 2017.
- (14) HAYKIN, Simon. Redes Neurais: princípios e práticas; trad. Paulo Martins Engel. – 2ª ed. – Porto Alegre. Bookman, 2001.
- (15) G1 MARANHÃO. Maranhão é o estado que mais registra fraudes na previdência social. São Luís, 2017. Disponível em: <https://g1.globo.com/ma/maranhao/noticia/maranhao-e-o-estado-que-mais-registra-fraudes-na-previdencia-social.ghtml> . Acesso em: 01 out. 2018.
- (16) G1 RIO GRANDE DO SUL. Quadrilha suspeita de fraude em salário-maternidade é alvo de operação da PF em Caxias do Sul. Porto Alegre, 2018. Disponível em: <https://g1.globo.com/rs/rio-grande-do-sul/noticia/quadrilha-suspeita-de-fraude-em-salario-maternidade-e-alvo-de-operacao-da-pf-em-caxias-do-sul.ghtml> . Acesso em: 05 out. 2018.
- (17) SECRETARIA DE PREVIDÊNCIA. Operação prende três pessoas por fraude em pensões por morte e salário-maternidade. Brasília, 2018. Disponível em: <http://www.previdencia.gov.br/2018/09/operacao-prende-tres-pessoas-por-fraude-em-pensoes-por-morte-e-salario-maternidade/> . Acesso em 10 de out. de 2018.
- (18) MINISTÉRIO DO DESENVOLVIMENTO SOCIAL E AGRÁRIO. Nota Técnica nº 7 de 2017 do Gabinete da Secretaria de Avaliação e Gestão da Informação do Ministério do Desenvolvimento Social e Agrário. Brasília, 2017.
- (19) NORMAS LEGAIS. Código de Benefícios da Previdência Social. Disponível em: <http://www.normaslegais.com.br/guia/clientes/codigo-beneficios-previdencia-social.htm> . Acesso em: 01 de nov de 2018.

- (20) VALENÇA, Mêuser. Fundamentos das Redes Neurais: exemplos em Java. Livro Rápido. Olinda, 2009.
- (21) LUDERMIR, T.B. et al. Redes neurais artificiais: teoria e aplicações. LTC Editora, Rio de Janeiro, 2007