

Universidade de Pernambuco - UPE
Campus Garanhuns
Curso de Licenciatura em Computação

**GUSTAVO HENRIQUE FERREIRA DE MIRANDA
OLIVEIRA**

**SISTEMA HÍBRIDO PARA PREVISÃO DE SÉRIES
TEMPORAIS EDUCACIONAIS**

Garanhuns – PE

Dezembro, 2015

**GUSTAVO HENRIQUE FERREIRA DE MIRANDA
OLIVEIRA**

**SISTEMA HÍBRIDO PARA PREVISÃO DE SÉRIES
TEMPORAIS EDUCACIONAIS**

Monografia apresentado ao Professor João Fausto Lorenzato de Oliveira, do Curso de Graduação em Licenciatura em Computação da Universidade de Pernambuco - Campus Garanhuns, em cumprimento às exigências para conclusão de curso.

ORIENTADOR: João Fausto Lorenzato de Oliveira

Garanhuns - PE

Dezembro, 2015

***Dedico ao meu Pai Gilberto Beserra e a minha mãe Maria
Luci.***

AGRADECIMENTOS

Primeiramente gostaria de agradecer a Deus por tudo que tem proporcionado na minha vida e por toda a força dada para conseguir passar pelos obstáculos que a vida impôs.

Agradecer a base da minha vida, meus pais Gilberto e Luci e meus irmãos Augusto e Marília, por todo carinho, assistência e confiança a mim dada.

Agradecer a todos os meus professores por todo conhecimento passado, pois foi de extrema importância para a minha formação profissional.

Agradecer ao meu Padrinho Acadêmico, professor Fausto Lorenzato pela orientação de todos os meus trabalhos, incentivo, paciência e motivação.

Agradecer a Hellen Carolina, por me acompanhar diariamente nesses 4 anos de graduação, por todo companheirismo nos momentos difíceis e felizes, a você lhe dedico um “cheiro” especial.

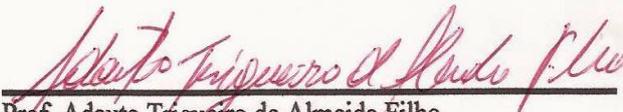
Agradecer ao meu irmão Augusto César por ser meu braço forte na graduação, por cada trabalho feito junto e por ser um irmão de verdade em todas as horas, lhe dedico um abraço forte!

Agradecer aos meus amigos de graduação, Michael (Nóia), Alisson, Yuri, Thiago, Rodrigo (Coroinha), Guilherme, Rafael, Ícaro, Ailson, Paulo Antônio e Jéssica que fizeram dos momentos chatos os mais engraçados possíveis.

Prepara-se o cavalo para o dia da batalha, porém do SENHOR vem a vitória.

Provérbios 21:31.

Monografia de Graduação apresentada por **Gustavo Henrique Ferreira de Miranda Oliveira** do Curso de Graduação de Licenciatura em Computação da Universidade de Pernambuco – Campus Garanhuns, sob o título “**Sistema híbrido para previsão de séries temporais educacionais**”, orientado (o) a pelo (a) Professor (a) **João Fausto Lorenzato de Oliveira** e aprovada pela Banca Examinadora formada pelos professores:


Prof. Adauto Trigueiro de Almeida Filho
Licenciatura em Computação / UPE


Prof. João Fausto Lorenzato de Oliveira
Licenciatura em Computação / UPE

Visto e permitida à impressão.
Garanhuns, 21 de Dezembro de 2015.


Prof. Emanuel Francisco Spósito Barreiros
Coordenador do Curso de Licenciatura em Computação
da Universidade de Pernambuco, Campi Garanhuns.

Emanuel Francisco Spósito Barreiros
Coordenador Curso de Computação
UPE
Emanuel Francisco Spósito Barreiros
Coordenador Curso de Computação
UPE Campus Garanhuns
Matricula: 12.088-0

RESUMO

A combinação de modelos de previsão tem sido uma estratégia utilizada na construção de sistemas híbridos que realizam previsão em séries temporais nas mais diversas áreas do conhecimento. O presente trabalho teve como objetivo a implementação de um modelo que combina a técnica de previsão linear, Autorregressivo (AR), a técnica de previsão não-linear Multilayer Perceptron (MLP) e o Algoritmo de Enxame de Partículas (PSO), para realizar previsões em séries temporais educacionais. Várias estratégias de combinação de previsões foram testadas visando aumentar a precisão das previsões, como: soma das previsões lineares e não-lineares, multiplicação das previsões por constantes lineares e a utilização do algoritmo BackPropagation. Para medir a qualidade das previsões geradas, foi utilizado a medida de desempenho de Erro Médio Quadrático (MSE). O modelo proposto realizou as melhores previsões em 6 dos 7 conjuntos de dados trabalhados.

Palavras-chave: Previsão, Séries Temporais Educacionais, Combinação de modelos de previsão.

ABSTRACT

The combination of forecast models has been a strategy used to build hybrid systems that perform time series forecasting in several areas of knowledge. This study aimed to implement a model that combines the linear prediction technique, Autoregressive (AR), the nonlinear prediction technique Multilayer Perceptron (MLP) and the Particle Swarm algorithm (PSO) to make predictions in educational time series. Various strategies of combining predictions were tested to increase forecast accuracy, as sum of linear and nonlinear predictions, forecasts by multiplying the linear constant and using the backpropagation algorithm. To measure the quality of the generated predictions, we used the average error performance measure Quadratic (MSE). The proposed model made the best forecasts in 6 of 7 sets of raw data.

Keywords: Forecasting, Time Series Educational, Combination forecasting models.

SUMÁRIO

| | |
|---|-----------|
| 1. Introdução | 13 |
| 1.1. Motivação | 15 |
| 1.2. Objetivos | 16 |
| 1.2.1. Objetivo geral..... | 16 |
| 1.2.2. Objetivos específicos | 16 |
| 1.3. Estrutura da Monografia | 17 |
| 2. Fundamentação Teórica | 18 |
| 2.1. Previsão de séries temporais. | 18 |
| 2.2. Algoritmo de Enxame de Partículas | 20 |
| 2.3. Modelo Autorregressivo | 21 |
| 2.4. Redes Neurais Artificiais | 22 |
| 3. Elaboração do método proposto | 24 |
| 3.1. Soma das Previsões | 28 |
| 3.2. Constantes de Linearidade | 29 |
| 3.3. Algoritmo de Retropropagação | 31 |
| 3.4. Vantagens e Desvantagens | 33 |
| 4. Experimentos | 35 |
| 4.1. Bases Utilizadas | 35 |
| 4.1.1. Séries educacionais do United States Census Bureau. | 35 |
| 4.1.2. Quantidade de estudantes que conseguiram concluir o ensino superior..... | 37 |
| 4.1.4. Número de manchas solares. | 38 |
| 4.1.5. Quantidade de veículos e peças produzidos no Canadá. | 39 |
| 4.1.6. Tabelas com abreviações das séries temporais. | 39 |
| 4.2. Configuração de experimentos..... | 40 |
| 4.3. Análise de resultados..... | 41 |
| 5. Conclusões e Trabalhos Futuros | 45 |
| 5.1. Conclusões | 45 |
| 5.2. Trabalhos Futuros | 46 |
| Referências | 47 |

Índices de Figuras

| | |
|--|----|
| Figura 1. Neurônio Artificial de McCulloch e Pitts..... | 22 |
| Figura 2. Estrutura do MLP..... | 26 |
| Figura 3. Estrutura da partícula. | 27 |
| Figura 4. Estrutura geral do sistema híbrido. | 28 |
| Figura 5. Estrutura do algoritmo com a soma das previsões. | 29 |
| Figura 6. Estrutura da partícula com as constantes..... | 30 |
| Figura 7. Estrutura do algoritmo com as constantes de linearidade. | 30 |
| Figura 8. Estrutura do algoritmo BP. | 31 |
| Figura 9. Estrutura do algoritmo geral com o algoritmo BP. | 32 |
| Figura 10. Base de entrada e saída do algoritmo BP..... | 33 |
| Figura 11. Pessoas latino-americanas que completaram o ensino fundamental em 4 anos ou menos. | 36 |
| Figura 12. Pessoas latino-americanas que completaram o ensino fundamental entre 5 a 8 anos. | 36 |
| Figura 13. Porcentagem de pessoas latino-americanas de 25 anos ou mais que concluíram a graduação. | 37 |
| Figura 14. Quantidade de concluintes de ensino superior na Austrália. | 37 |
| Figura 15. Quantidade de passageiros das companhias aéreas internacionais. | 38 |
| Figura 16. Número de manchas solares entre 1700 a 1987..... | 39 |
| Figura 17. Quantidade de veículos e peças vendidos no Canadá. | 39 |
| Figura 18. Fluxo do experimento. | 40 |
| Figura 19. Previsão do modelo proposto em comparação com a série Conclusão 4 anos..... | 42 |
| Figura 20. Previsão do modelo proposto em comparação com a série Conclusão 5 a 8 anos. | 43 |
| Figura 21. Previsão do modelo proposto em comparação com a série Conclusão Graduação..... | 43 |
| Figura 22. Previsão do modelo proposto em comparação com a série Conclusão Austrália..... | 43 |
| Figura 23. Previsão do modelo proposto em comparação com a série Aéreas..... | 44 |
| Figura 24. Previsão do modelo proposto em comparação com a série Manchas. | 44 |

Figura 25. Previsão do modelo proposto em comparação com a série Veículos.
.....44

Índices de Tabelas

| | |
|---|----|
| Tabela 1. Entradas e saídas desejadas do MLP com ordem três. | 26 |
| Tabela 2. Soma dos elementos das previsões do AR e MLP. | 29 |
| Tabela 3. Multiplicação das constantes nas previsões do AR e MLP..... | 30 |
| Tabela 4. Resumo das características das séries temporais utilizadas. | 40 |
| Tabela 5. Tabela de Experimentos..... | 42 |

1. Introdução

É importante que toda e qualquer instituição educacional busque avanços para o seu progresso e para isso é necessário sempre repensar o seu papel como instituição sabendo quais metas deseja atingir, atualizando-se e inovando-se sempre. Muitos são os problemas de âmbito nacional que permeiam as instituições educacionais, por exemplo, o mau rendimento escolar, repetência e evasão. Cada instituição é responsável por criar ações para combater estes problemas, logo é imprescindível conhecer estes índices, pois somente assim é possível criar ações de combate a fim de minimizá-los.

A educação é uma área de forte influência nacional, pois qualquer mudança realizada nesse âmbito tem o poder de afetar diretamente o desenvolvimento do nosso país. Considerar índices educacionais e desenvolver mecanismos que ofereçam suporte nesta área tão importante é indispensável, pois torna viável o acompanhamento do desenvolvimento social, econômico e demográfico do país.

É importante analisar e estimar esses índices para saber os motivos deles surgirem, para que de acordo com as demandas existentes possibilite um maior controle das fases de manutenção, tendo em vista a redução e execução dos custos para melhorá-los, permitindo a ampliação e aprofundamento das estratégias de apoio.

A coleção destes índices observados ao longo de um determinado intervalo de tempo pode ser considerada uma série temporal que é um conjunto de dados de uma mesma variável ou área pesquisada. Segundo Moretin (1987) uma série temporal “é um conjunto de observações ordenadas no tempo, não necessariamente igualmente espaçadas, que apresentam dependência serial, isto é, dependência entre instantes de tempo”.

Conforme Moretin (1987), séries temporais possuem comportamentos distintos, esses tipos de comportamentos podem ser classificados em quatro elementos:

1. Tendência: onde se verifica o sentido de deslocamento da série no decorrer de vários anos.
2. Ciclo: a série temporal possui movimento ondulatório que ao longo de vários anos tende a ser periódico.
3. Sazonalidade: a série temporal possui movimento ondulatório de curta duração, e no geral é inferior a um ano, onde muitas vezes é associado a mudanças climáticas.
4. Ruído aleatório ou erro: que compreende a variabilidade dos dados e não pode ser modelado.

No mundo real os dados de uma série temporal quase nunca são inteiramente lineares, ou seja, que possuem padrões lineares. De acordo com Zhang (2001) “séries temporais podem ser consideradas como a composição de dois padrões, os lineares e não-lineares”. Este tipo de formação de dados impede o entendimento e leitura, pois quase sempre não se sabe a relação entre eles, dificultando a tomada de decisões.

Portanto reconhecer os padrões que as séries temporais possuem torna-se uma tarefa bastante complexa, pois no mundo real os dados são estocásticos, alterando-se com o decorrer do tempo.

Alguns estudos propostos na literatura como o de Yolcu (2012) tem o objetivo de prever os dados futuros das séries temporais. Para a realização desta tarefa o autor propõe a utilização de técnicas de Inteligência Artificial (IA) a fim de utilizá-las para realizar previsões. No entanto segundo Zhang (2001) nenhum modelo individual de previsão é melhor em todas as situações. Devido a isso vários estudos empíricos como o de Oliveira (2014) e Robinson (2001) sugerem que a combinação de vários modelos diferentes possam melhorar as previsões em relação a um modelo único e individual.

Na literatura, existem diversas abordagens híbridas, ou seja, sistemas que fazem uso da combinação de modelos de previsões. Nos métodos propostos geralmente os autores fazem uso de duas fases de previsões para lidar com as séries temporais, dentre estes trabalhos destaca-se o de Adhikari (2015) e Lee (2012). Na primeira fase os componentes lineares da série são modelados por um modelo linear e na segunda fase os resíduos deixados pela primeira modelagem são tratados por um método não-linear, presumindo que nessa fase só existam componentes não-lineares, pois os lineares teriam sido extraídos anteriormente.

Seguindo esta organização de fases encontrada na literatura para modelar as séries temporais, as Redes Neurais Artificiais (RNA) são utilizadas neste trabalho com o intuito de fazer as previsões não-lineares, pois segundo Portugal (1996) as RNA tem uma melhor capacidade de prever estes tipos de dados.

A combinação de modelos como o Autorregressivo (AR) que segundo Zhang (2001) possui capacidade de modelar dados lineares, juntamente com as Redes Neurais (RNA) pode melhorar a precisão da previsão, devido a cada técnica desempenhar uma melhor previsão para cada tipo de dados.

Outro fator que influencia a utilização destas técnicas é a inicialização dos pesos que as técnicas de previsão irão utilizar. Para um dado problema, diferentes inicializações dos pesos podem fornecer diferentes resultados. A escolha destes pesos pode ser feita de modo tentativa erro, porém não trará um bom resultado além de consumir muito tempo. Sabendo-se disto esse tipo de

problema passa a ser de otimização, onde é necessário fazer uma busca pelo melhor conjunto de pesos em um espaço dimensional encontrando uma solução ótima.

Segundo Oliveira (2014) sistemas evolutivos como o algoritmo de enxame de partículas (PSO), tem como base de inspiração o voo das aves. E cada indivíduo ou partícula representa uma solução potencial para a otimização do problema, realizando movimentos com base na melhor solução do enxame, ou seja, com base na melhor partícula. Este método também tem a capacidade de otimizar os parâmetros de outras técnicas de previsão, podendo assim otimizar os pesos da função do AR e RNA para que os mesmos tenham melhores resultados nas suas previsões.

1.1. Motivação

A previsão de séries temporais é uma área dinâmica de pesquisa contendo uma importância fundamental em numerosos campos práticos de atuação. Melhorar a precisão das séries temporais é uma tarefa extremamente desafiadora, pois reconhecer padrões nos dados da vida real em que eles estão em constante mudança requer um grandioso trabalho. Segundo Adhikari (2015) a previsão de séries temporais vem ganhando atenção e pesquisa contínua em poucas décadas, devido a necessidade de obter estes dados nos mais diversos problemas.

Estes dados das séries temporais são bastante importantes, pois a partir destes podemos desenvolver um modelo que poderá descrever uma relação subjacente, ou seja, poderá prever uma série temporal futura. A previsão de séries temporais ajuda tanto corporações educativas como também corporações empresariais. Segundo Oliveira (2014), a previsão de séries temporais “pode ser beneficiada pelo uso de sistemas evolutivos que melhoram a previsão destes modelos e causam impacto positivo sobre o processo de tomada de decisões de muitas organizações”.

O Índice de Desenvolvimento da Educação Básica (Ideb) é um índice muito importante para a educação no Brasil e é calculado sobre dois conceitos: fluxo escolar e média de desempenho nas avaliações: Inep, Saeb e Prova Brasil. Gerado anualmente pelo Censo Escolar com a intenção de comparar estatisticamente os dados do mesmo ano e de anos diferentes, permitindo assim traçar metas de qualidade educacional para todos os sistemas envolvidos.

Segundo o Portal Inep, o Ideb trata-se do mais relevante e abrangente levantamento estatístico sobre a educação básica do Brasil. Os dados coletados constituem uma fonte completa de informações utilizada pelo Ministério da Educação (MEC) para formular políticas e para desenhar programas, bem como para a definir critérios para a atuação supletiva do MEC às escolas, aos estados e aos municípios.

Vários índices são analisados para melhorar a educação no país o Ideb é um deles provando a importância dos índices educacionais para elaboração de metas. Portanto prever dados relacionados à educação tais como reprovação, repetência e evasão escolar podem fazer com que os mesmos levem discussões para os dirigentes, coordenadores de cursos e docentes, a fim de tomar ações firmes e efetivas para combatê-los.

Em suma a importância das séries temporais deve-se ao fato de poder retratar e acompanhar os problemas que assolam a sociedade como um todo. Na literatura existem vários estudos que buscam através dos sistemas híbridos, prever dados em diferentes tipos de séries temporais, pois como nos exemplos citados a previsão destes dados afeta diretamente no planejamento das instituições.

Deste modo a importância das séries temporais é notável, pois com estas é possível fazer um levantamento de todos os fatores que estão afetando a área pesquisada, permitindo a criação de modelos de previsões que busquem oferecer dados confiáveis para a tomada de decisões de todas as pessoas envolvidas.

Logo desenvolver ações em cima dos índices futuros será mais uma ferramenta para melhorar a educação brasileira, de modo que decisões importantes sejam tomadas criando programas de permanência de alunos e métodos que melhorem os índices de repetência e reprovação.

1.2. Objetivos

Nesta seção os objetivos do trabalho serão descritos.

1.2.1. Objetivo geral

Elaborar um sistema híbrido e evolutivo que combina um conjunto de técnicas diferentes de IA com o intuito de realizar previsões mais precisas. Utilizando o algoritmo de otimização de enxame de partículas (PSO) para definição dos parâmetros da função das RNA para os dados não-lineares e para o AR nos dados lineares.

1.2.2. Objetivos específicos

- Elaborar um estudo comparativo sobre a combinação dos modelos AR e MLP treinados com métodos de otimização PSO em relação a técnicas convencionais como o BackPropagation (BP) e AR e MLP, submetidos à otimização.
- Apresentar uma nova metodologia de combinação de preditores, fornecendo assim previsões mais eficientes.

- Generalizar a metodologia de forma a tratar séries temporais que apresentam características diversas testando em séries reais encontradas da literatura.
- Aplicar o algoritmo proposto á séries educacionais realizando previsões sobre dados educacionais de reprovação, repetência e evasão, com a intenção de gerar discussões para os dirigentes.

1.3. Estrutura da Monografia

Este trabalho encontra-se estruturado em 5 capítulos. O capítulo 2 apresenta a fundamentação teórica, o capítulo 3 descreve o Método Proposto. O capítulo 4 os Experimentos realizados, comparações entre métodos e resultados, por fim o capítulo 5 com as conclusões e trabalhos futuros.

2. Fundamentação Teórica

A revisão da literatura será feita nesta seção.

2.1. Previsão de séries temporais.

Uma série temporal é um conjunto de valores numéricos que foram observados e guardados durante determinado tempo, seja com intervalos de horas, dias, meses, trimestres ou anos. As observações existentes nas séries temporais independem de área, variando em diversos problemas.

A previsão dos dados das séries temporais é imprescindível, pois com eles a população ou os responsáveis podem se programar. Dependendo do tipo de dado que esteja sendo trabalhado o planejamento em cima deles pode variar, pois se uma empresa visa aumentar suas vendas tomarão como base estes dados para criar meios para ascendê-los, mas se forem dados que prejudiquem a empresa a intenção é diminuí-los.

Como as séries estão em todos os âmbitos e áreas, podemos citar alguns tipos de séries como: quantitativo de alunos que concluíram o ensino médio anualmente; previsão de catástrofes naturais como terremotos; quantidade de passageiros mensalmente em uma empresa aérea; valores mensais de temperatura registrados em uma cidade; valores diários de poluição em uma cidade e; índices diários da bolsa de valores.

Segundo Werner (2006) realizar previsões de demanda é uma atividade muito importante, pois pode revelar tendências de mercado e contribuir no planejamento estratégico da empresa. Realizar previsões e obter dados do futuro é de extrema importância, pois somente assim poderemos criar estratégias para contornar as situações problemáticas como também traçar metas para alcançar casos de sucesso, beneficiando empresas privadas, públicas e a sociedade em geral.

Segundo Zhang (2001) séries temporais é uma área de pesquisa ativa nas últimas décadas devido ao impacto gerado a sociedade envolvida. Entre os trabalhos encontrados na literatura que estão relacionados com séries temporais e a sua previsão destaca-se o de Reyes (2012) que utilizou como principal motivação do seu trabalho desenvolver um sistema que fosse capaz de prever dados de terremotos em quatro áreas sismogênicas do Chile, a fim de permitir que as autoridades planejem políticas preventivas, como desligar sistemas de estações de energia, transporte, telecomunicações e também reduzir o número de mortes de pessoas. O autor utilizou um algoritmo Feedforward com três camadas escondidas na RNA.

Cadenas (2010) teve como motivação do seu trabalho a necessidade que empresas eólicas como o Centro de Energia e Controle (CEC) tinham para fornecer eletricidade a curto e médio prazo. A intermitência e aleatoriedade dos

ventos era o principal problema, para isto o autor utilizou um modelo híbrido que faz a relação entre os modelos de previsão Autorregressivo Integrado de Movimento Média Móvel (ARIMA) e RNA e utilizou como séries temporais as observações dos dados de três regiões distintas no México.

Faruk (2009) propôs uma abordagem de modelagem de séries temporais para a qualidade de água do Rio Menderes, pois a economia agrícola da bacia do rio depende da água para irrigada, cultivo de algodão, milho, figueiras e oliveiras. A deterioração da qualidade da água desse rio resultou em impactos negativos para peixes, animais selvagens e vida aquática, levando a mortandade destes e considerações desagradáveis para o turismo. O autor utilizou um modelo ARIMA e RNA como abordagem, e o seu modelo deu previsões confiáveis sobre a temperatura da água, boro e dados de séries temporais de oxigênio dissolvido.

Extensos trabalhos como Oliveira (2014), Adhikari (2015), Tseng (2000) e Lee (2012) foram realizados com o intuito de combinar vários modelos de previsões de séries temporais com a conclusão geral de que esta prática melhora a previsão e exatidão em grande medida. Além disso, sistemas inteligentes que combinam técnicas de previsão, frequentemente batem todos os modelos individuais em termos de precisão de previsão.

Nestes trabalhos destaca-se o de Oliveira (2014) que buscou desenvolver um sistema híbrido que fosse robusto e eficiente em séries temporais distintas, utilizando o algoritmo PSO e as técnicas de previsão ARIMA e Máquinas Vetores de Suporte (SVR). As séries utilizadas foram: Passageiros da empresa internacional Airlines; Estoque da empresa IBM; Preço da ação comum anual; Armadilhas de lincas no canada; Demanda mensal por gasolina; Vendas para impressão e escrita e; Número mensal de suínos abatidos.

Zhang (2001) propôs utilizar as técnicas ARIMA e RNA para modelagem linear e não linear das séries: Manchas solares; Quantidade de lincas canadenses e; Taxa de câmbio semanal da libra britânica com o dólar. Os resultados obtidos sugeriram claramente que o modelo é capaz de realizar uma boa previsão em cada série adotada.

Yolcu (2012) propôs um modelo de RNA que modela tanto dados lineares como não lineares ao mesmo tempo e por meio do algoritmo PSO otimizou seus pesos. A aplicação do seu modelo se destinou a três séries temporais, a primeira sobre a quantidade de dióxido de carbono medido mensalmente em Ancara capital da Turquia, a segunda relacionada ao consumo trimestral de cerveja australiana, e por ultimo a armadilhas de lincas no Rio Mackenzie no Canada.

Adhikari (2014) propôs uma abordagem com RNA de previsão combinada que determina os pesos e reconhece a relação entre os grupos destes em amostra. Foi feito um estudo com quatro modelos de previsões e oito séries

temporais, tais séries foram: Quantidade anual de lince presas no Canadá; Número anual de manchas solares observadas; Fluxo mensal do rio Clearwater em Kamiah, Idaho; Venda mensal de veículos nos EUA; Valores reais e anuais do Produto Nacional Bruto dos EUA; Venda mensal de vinho australiano; Linha aérea contendo o número mensal de passageiros das companhias aéreas internacionais e; Observações industriais trimestrais.

Lee (2010) utilizou como motivação do seu trabalho a busca por um sistema híbrido que realizasse previsões precisas em diferentes tipos de séries temporais. Foi desenvolvido um sistema que combina a relação dos modelos ARIMA e Programação Genética (GP), o modelo foi testado nas séries temporais de lince canadenses; Consumo de energia da China e; Produto interno bruto trimestral dos EUA.

2.2. Algoritmo de Enxame de Partículas

Segundo Prata (2012), na década de 90 cientistas das áreas de computação e zoologia passaram a investigar o comportamento psicológico e social de determinadas espécies, visando entender como as populações destes seres reagem com as adversidades do cotidiano, desde procurar por comida a fugir de predadores.

O estudo do comportamento social destes indivíduos motivou aos pesquisadores da época a simular estes por intermédio de computadores acreditando que esse procedimento ajudaria na solução de problemas computacionais. Kennedy e Eberhart (1995) fascinados pelo comportamento do voo dos pássaros propuseram uma nova meta-heurística destinada à resolução de problemas irrestritos de otimização não-linear, denominada Particle Swarm Optimization (PSO).

O PSO possui uma população de indivíduos denominada de enxame, onde cada indivíduo ou partícula é caracterizado por sua aptidão, obtida da sua função de avaliação. Cada partícula representa uma posição no espaço de busca e uma velocidade que regula essas posições.

Souza (2011) em seu trabalho explica a filosofia básica do algoritmo da seguinte forma: O PSO inicia seu processo de busca com um enxame aleatório de partículas. Cada partícula representa uma solução individual para o problema a ser resolvido, tendo como as principais etapas do algoritmo:

1. X_i , posição do agente a_i no espaço de soluções, onde cada partícula representa uma solução individual;
2. V_i , velocidade do agente a_i indicando a direção que a partícula deve assumir no espaço de busca;
3. pBest é a melhor posição encontrada por cada partícula individualmente.

4. gBest é a melhor posição global encontrado pelas partículas vizinhas.

Para cada iteração, as partículas movem-se em um espaço dimensional de tamanho equivalente a sua quantidade, sendo influenciadas pela sua própria experiência o pBest, e pela experiência da melhor solução do enxame o gBest. As partículas mudam de posição e velocidade de forma contínua a cada iteração, com o objetivo de alcançar melhores posições e melhorar a função de avaliação.

Alguns trabalhos como o de Oliveira (2014) e Souza (2011), utilizaram a força de otimização do PSO para achar a melhor solução para os seus problemas. Oliveira (2014) utilizou o PSO para encontrar o melhor conjunto de parâmetros para o ARIMA e SVR, métodos de previsão de séries temporais, com o intuito de aumentar a capacidade de previsão do sistema.

Souza (2011) investigou o uso do PSO para seleção de vários objetivos de Casos de Teste (TC), onde o algoritmo provê para o usuário um conjunto de soluções de teste em diferentes valores de cobertura, requisitos contra esforço e execução, onde o usuário pode então escolher a solução que melhor se encaixa aos seus recursos disponíveis.

2.3. Modelo Autorregressivo

O modelo autorregressivo ou comumente chamado de Box-Jenkins tem esse nome devido aos seus criadores os estatísticos George Box e Gwilym Jenkins. Este modelo foi sistematizado em 1976 sendo muito utilizado na modelagem e previsão de séries temporais.

Segundo Andrade (2011) essa metodologia permite a previsão de valores futuros de uma série temporal tomando como base a correlação temporal dos valores passados e presentes da série. A estrutura do AR é expressa por:

$$Z_t = \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \dots + \theta_p Z_{t-p} \quad (2)$$

Onde θ_i são os parâmetros da estrutura, produzindo um resíduo, $E_t = Z_t - Z_t'$ (3) onde Z_t' é a estimativa do componente linear.

O modelo autorregressivo estima os componentes lineares da série temporal, através de uma função com p observações passadas, $Z_{t-1}, Z_{t-2}, Z_{t-3}, \dots, Z_{t-p}$, onde p determina o número de passos entre as observações passadas e a previsão da próxima observação.

Para analisar tendências de séries temporais Latorre (2001) conta que é possível utilizar tanto funções matemáticas como modelos de regressão polinomial baseados na série inteira ou em determinado ponto da mesma. Mas o modelo autorregressivos possui uma capacidade maior, na qual tem a habilidade de fazer essa análise tanto em séries de comportamento de tendência como também sazonalidade.

Sabendo-se que as séries temporais possuem dados sequenciais coletados ao longo dos anos o modelo autorregressivo tem como principal objetivo modelar e captar os padrões lineares da série temporal, buscando realizar previsões futuras.

2.4. Redes Neurais Artificiais

As RNA fazem parte de uma vasta quantidade de técnicas existentes na IA, onde as mesmas possuem a capacidade de realizar operações de aprendizado, associação, generalização e abstração. Essas técnicas foram inspiradas sobre estudos da estrutura do “cérebro humano para tentar emular sua forma inteligente de processar informação”, (Portugal, 1996).

Segundo Fernandes (1995) as RNA funcionam conceitualmente de forma similar ao cérebro humano, tentando reconhecer regularidade entre padrões de dados, sendo capazes de aprender com a experiência e fazer generalizações baseadas no seu conhecimento previamente acumulado.

As RNA possuem como características alguns elementos que dão suporte a sua estrutura, tais elementos são os neurônios artificiais. Os neurônios estão interconectados e efetuam operações simples, transmitindo seus resultados aos neurônios vizinhos, tendo a habilidade de realizar mapeamentos não-lineares entre suas entradas e saídas tornando esta estrutura ótima para o reconhecimento de padrões.

A Figura 1 representa o neurônio artificial proposto pelo psiquiatra e neurofisiologista McCulloch e o matemático Pitts, possuindo as seguintes características:

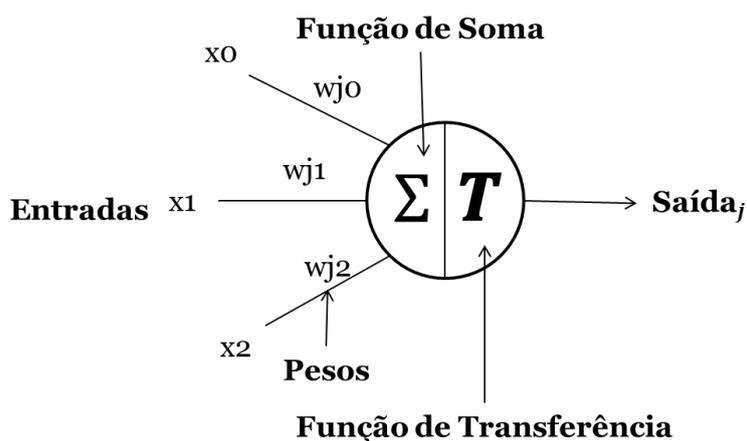


Figura 1. Neurônio Artificial de McCulloch e Pitts.

Cada componente do neurônio artificial corresponde a um elemento do neurônio biológico para simular os comportamentos reais. Onde **X0**, **X1** e **X2** representam as entradas que equivalem aos dendritos, à saída equivalendo ao axônio e para simular a sinapse cada conexão do neurônio possui um peso

sendo W_{j0} , W_{j1} e W_{j2} . Esses pesos contribuem para o resultado de saída do neurônio determinando com que intensidade o sinal chegará. O corpo celular do neurônio biológico é simulado pela soma dos valores do produto de suas entradas pelos seus respectivos pesos, sendo representado pela função soma Σ , sua ativação e o seu limiar (threshold) é definido pela função de transferência T ou função de ativação.

Essa ativação do neurônio artificial tem como finalidade restringir a intensidade de saída do neurônio, tendo seu valores normalizados no intervalo de (0, 1) ou (-1, 1).

Na literatura, encontram-se muitos tipos de redes neurais com as mais diversas arquiteturas e algoritmos de aprendizado. Ainda que existam diversas quantidades de arquiteturas, uma das mais utilizadas e a estrutura multicamada (MLP). Esta estrutura é uma das mais utilizadas devido a sua capacidade de aproximação universal.

Um algoritmo de aprendizado muito utilizado é o Backpropagation (BP) que tem como características utilizar um conjunto de entradas e saídas para realizar o treinamento das estruturas para efetuar o aprendizado supervisionado. A partir do conjunto de treinamento a rede realiza o seu processamento obtendo uma saída. Esta saída é comparada com uma saída esperada, no intuito de gerar um conjunto de erros. A partir destes erros o processo de ajustes de pesos é aplicado, buscando minimizar estes conjuntos.

Segundo Fernandes (1995) RNA tem se mostrado extremamente útil nos mais diversos problemas, pois elas possuem uma a capacidade de fazer aproximações universais, tanto em reconhecimento de padrões, classificações e previsões.

Como o processo de previsão de séries temporais está diretamente ligado aos dados e eventos passados e no reconhecimento do comportamentos das séries temporais as RNA podem trabalhar bem neste tipo de problema possuindo a capacidade de aprender com os dados passados e por fim realizar previsões.

3. Elaboração do método proposto

O método proposto tem como finalidade realizar previsões em séries temporais com as mais diversas arquiteturas e tendências, possuindo robustez suficiente para mudança dos dados.

Para concebermos a estrutura do método proposto é necessário entender que segundo Zhang (2001) série temporal é uma composição de diferentes tipos de dados tendo uma estrutura de autocorrelação linear e de componentes não lineares, sendo expresso por,

$$y_t = L_t + N_t \quad (1)$$

Onde L_t indica os componentes lineares da série e N_t indica os componentes não-lineares. Muitos autores na literatura, como Oliveira (2014) fazem a combinação de diferentes modelos de previsões como estratégia para conseguir lidar com essa composição das séries temporais.

Nos métodos propostos as abordagens geralmente consistem em duas fases, onde na primeira é utilizado uma técnica de modelagem linear e na segunda uma técnica de modelagem não-linear para modelar os resíduos deixados da primeira fase. Nessas estratégias de duas fases presume-se que na primeira fase só existam componentes lineares e na segunda uma estrutura não-linear.

Existem diversos modelos de previsão de séries temporais, tanto lineares como também não-lineares, Ribeiro (2011) aponta alguns desses modelos, entre os modelos lineares é possível destacar os seguintes: AR, ARIMA, Média Móvel (MA) e Autorregressivo Condicional de Heteroscedasticidade (ARCH). Como modelos de previsão não-lineares as RNA e SVR.

Este trabalho propõe uma abordagem híbrida de duas fases que consiste em combinar dois modelos de previsões de séries temporais o AR para mapear os dados lineares e as RNA para fazer os mapeamentos não-lineares. Os pesos utilizados nas duas técnicas são encontrados por meio de otimização realizados pelo algoritmo PSO. Ao final são utilizadas três abordagens distintas para combinar e achar relações entre as previsões dos dois modelos.

O primeiro passo do método proposto consiste em estimar os componentes lineares através do modelo AR. Esse modelo tem como característica utilizar uma função com p de observações passadas da série temporal, como: $Z_{t-1}, Z_{t-2}, Z_{t-3}, \dots, Z_{t-p}$, onde p é o numero de passos entre as observações passadas e a previsão da próxima observação, sendo calculado como a equação 2:

$$Z_t = \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \dots + \theta_p Z_{t-p} \quad (2)$$

θ_i são os parâmetros da estrutura. Essa computação tem como finalidade produzir um resíduo, $E_t = Z_t - Z_t'$ (3) onde Z_t' é a estimativa do componente linear.

Como a maioria das séries temporais quase sempre contem padrões não-lineares, somente modelos lineares não podem produzir resultados satisfatórios quando tais séries temporais são analisadas. Para prever este tipo de dados nas abordagens de duas fases, modelos não-lineares são alimentados com os resíduos. Segundo Yolcu (2011) RNA tem sido muito utilizado na literatura. No entanto, modelos não-lineares não são suficientes por si só nas previsões.

Na segunda fase da abordagem, utilizamos a investida de fase única das RNA para modelar os resíduos do AR, utilizando a arquitetura MLP. Segundo Barros (2005) as redes MLP não foram concebidas para tratar problemas de séries temporais, devido a isto o método janela de tempo foi desenvolvido para criar uma adaptação da rede MLP para o processamento dinâmico.

Quando utilizamos redes MLP para previsão de séries temporais, definimos como entrada um conjunto de valores passados ordenados no tempo e como saída o valor que precede a esta sequência. Espera-se que com os padrões de entrada no treinamento consiga-se achar uma relação com a saída esperada.

Para tanto é necessário realizar um processamento na série original, que é feito através da equação:

$$E'_{(t)} = f(E_{t-1}, E_{t-2}, E_{t-3} \dots, E_{t-p}) \quad (4)$$

Onde $E_{(t)}$ é o ponto da série correspondente ao tempo t, p é o tamanho da janela de tempo e $E'_{(t)}$ é a nova série gerada. Devido a essa adaptação a série gerada terá pontos a menos que a série original, mas nenhum valor é omitido.

Para entendermos melhor como se comporta a janela de tempo, tomemos como base a série $E_{(1)}, E_{(2)}, E_{(3)}, E_{(4)}, E_{(5)}, \dots, E_{(n)}$, com a janela variando de 1 a n. Tomando como exemplo a série gerada com ordem igual a 3, utilizaremos os três elementos iniciais da série para serem nossos padrões de entrada, e por conseguinte a saída desejada seria o valor que o precede, ou seja, as entradas seriam $E_{(1)}, E_{(2)}, E_{(3)}$ e a saída esperada $E_{(4)}$.

Todas as saídas da série são analisadas por meio de varias iterações sempre seguindo essa mesma ordem, na segunda iteração deste exemplo as entradas seriam $E_{(2)}, E_{(3)}, E_{(4)}$ e a saída desejada $E_{(5)}$ e assim sucessivamente para o restante dos pontos da série, como exemplo na Tabela 1.

| Entradas | Saídas |
|-----------------------------------|------------|
| $E_{(1)}, E_{(2)}, E_{(3)}$ | $E'_{(4)}$ |
| $E_{(2)}, E_{(3)}, E_{(4)}$ | $E'_{(5)}$ |
| $E_{(3)}, E_{(4)}, E_{(5)}$ | $E'_{(6)}$ |
| $E_{(4)}, E_{(5)}, E_{(6)}$ | $E'_{(7)}$ |
| ... | ... |
| $E_{(n-3)}, E_{(n-2)}, E_{(n-1)}$ | $E'_{(n)}$ |

Tabela 1. Entradas e saídas desejadas do MLP com ordem três.

A estrutura do MLP utilizado neste trabalho é ilustrada na Figura 1, e a função utilizada é a função sigmoide onde f_1 é dada por:

$$f_1(x) = \frac{1}{1 + e^{-x}} \quad (5)$$

Como pode ser visto na Figura 1 foi utilizada somente uma camada de neurônios escondidos, w_1 é um vetor que inclui os pesos entre as entradas e os neurônios escondidos, w_2 inclui os pesos entre a camada escondida e a saída esperada. Como pode ser analisado na Figura 2, na saída não utilizamos a função sigmoide, pois nosso problema é direcionado para previsão de séries temporais e não para classificações de dados.

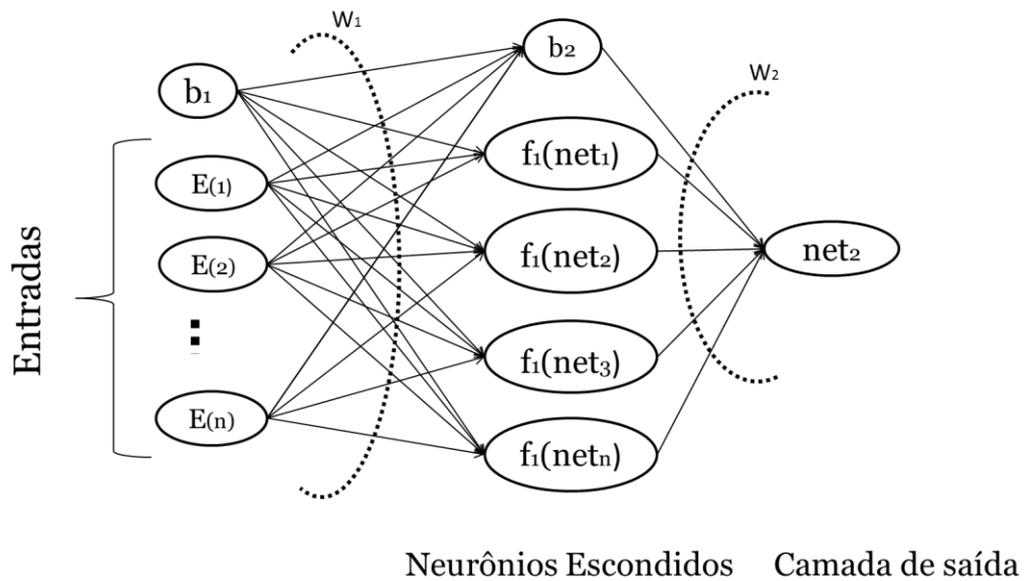


Figura 2. Estrutura do MLP.

Para obter as saídas esperadas o processo do MLP é dado por dois estágios:

Estagio 1: É calculado a saída do neurônio da camada escondida a partir da sua entrada correspondente. O valor de ativação para cada neurônio é obtido através dos net_n dada da seguinte formula:

$$net_n = [\sum_{j=1}^m w_{1j}E_{t-j} + b_1] \quad (6)$$

Onde w_{1j} ($j = 1, 2, 3, \dots, m$) são elementos de w_1 e b_1 é o bias. A função de ativação utilizado no neurônio é f_1 dada na Equação 5. De modo que o valor de saída o_n é calculado desse modo:

$$o_n = f_1(net_{1n}) \quad (7)$$

Estagio 2: Após o cálculo de todas as funções dos neurônios escondidos, o valor do neurônio de saída é calculado:

$$net_2 = [\sum_{j=1}^m w_{2j}net_{1j} + \dots w_{2m}net_{1m} \dots + b_2] \quad (8)$$

Neste estudo os pesos são treinados usando o algoritmo PSO tradicional devido a sua grande capacidade de busca no espaço dimensional. A busca no espaço é limitada por valores mínimo e máximo atribuídos para os parâmetros.

Nas posições do PSO todas as partículas contêm tanto os pesos do AR como do MLP e a quantidade de posições que cada partícula terá é determinada pela seguinte formula:

$$D = p + 1 + (p * N) + (2 * N) + 1 \quad (9)$$

Onde p é a ordem utilizada no AR e N é a quantidade de neurônios que são utilizados no MLP. A estrutura de uma partícula é ilustrada na Figura 3.

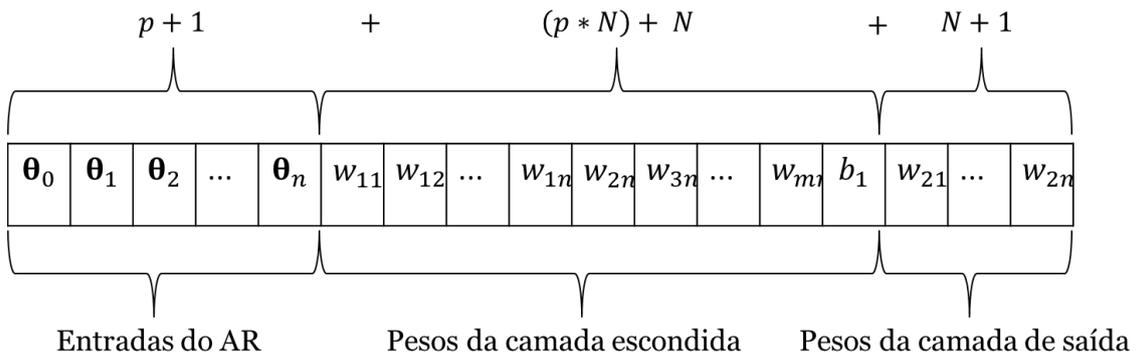


Figura 3. Estrutura da partícula.

O PSO realiza a pesquisa no espaço dimensional referente ao tamanho da formula (9) no qual é dado antecipadamente. O fluxo geral de execução do algoritmo deste trabalho é mostrado na Figura 4.

| | | |
|-------|-------|-------------------|
| E_5 | C_5 | $A_5 = E_5 + C_5$ |
| ... | ... | ... |
| E_n | C_n | $A_n = E_n + C_n$ |

Tabela 2. Soma dos elementos das previsões do AR e MLP.

A avaliação da aptidão do algoritmo PSO baseia-se no Erro Quadrático Médio (MSE) da previsão final do sistema. O MSE foi utilizado para avaliar a qualidade da previsão final. Sua respectiva formula é dada na Equação 10.

$$MSE = \frac{1}{n} \sum_{t=1}^n (Saída_t - Alvo_t)^2 \quad (10)$$

$Saída_t$ seria a saída esperada e $Alvo_t$ seria o elemento da previsão final que corresponde a saída desejada.

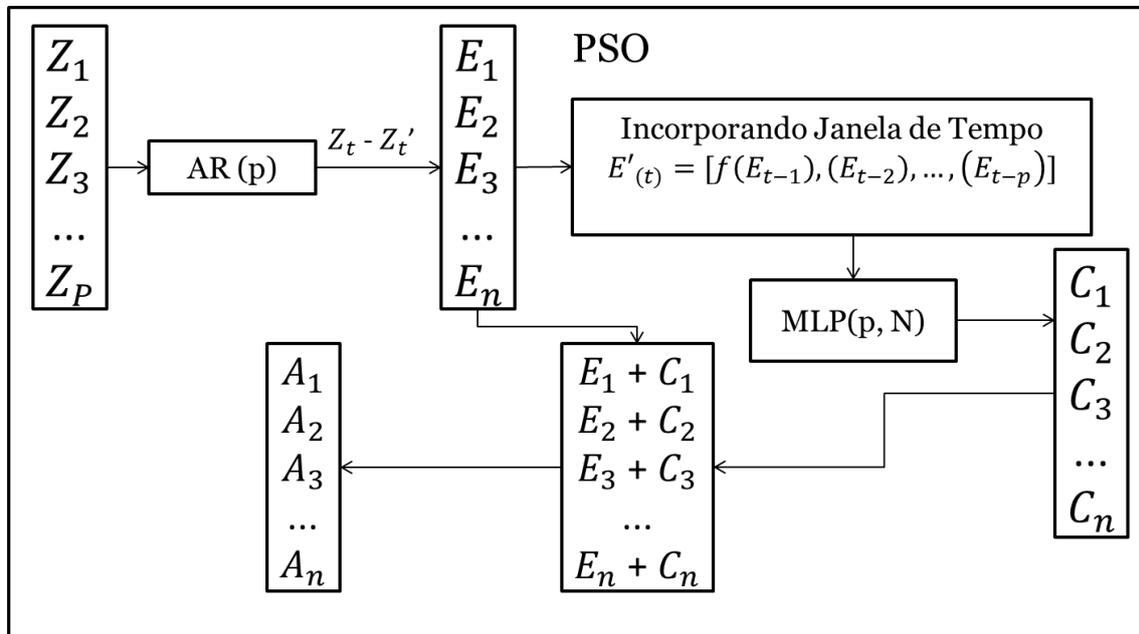


Figura 5. Estrutura do algoritmo com a soma das previsões.

3.2. Constantes de Linearidade

Nas abordagens de duas fases presumimos que todos os dados tenham sido capturados obtendo assim uma previsão ótima. Porém como as previsões estão ligadas diretamente aos pesos encontrados, nem sempre em todas as execuções do algoritmo geral se terá a melhor solução possível, desse modo não podemos afirmar que todos os dados foram capturados com êxito.

Nesta abordagem subtende-se que as duas previsões iniciais possuem uma relação linear, e essa relação pode melhorar as previsões. Na construção das partículas do PSO foram adicionadas duas constantes, uma para cada previsão como segue na Figura 6. Essas constantes serão multiplicadas por cada

elemento das duas previsões para obtermos uma previsão final, como segue a Tabela 3.

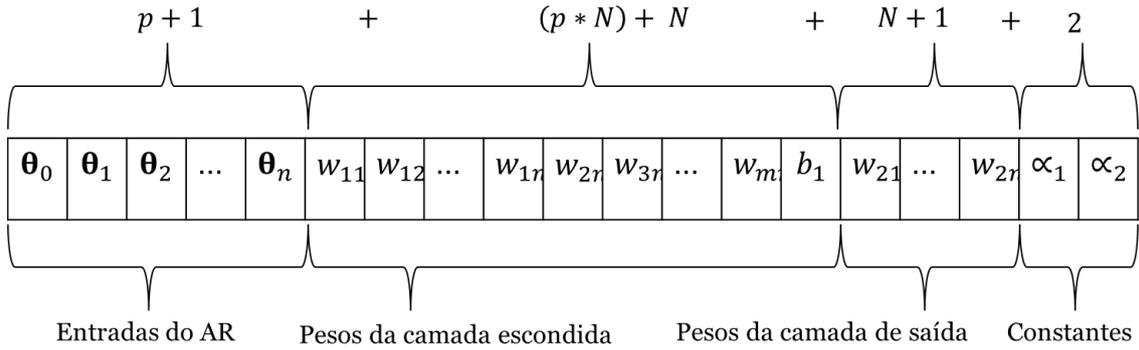


Figura 6. Estrutura da partícula com as constantes.

A avaliação de aptidão do algoritmo PSO nesta abordagem também utilizou o MSE descrito na Equação 10, para avaliar a qualidade da previsão final.

| Previsão AR | Previsão MLP | Previsão Final |
|-------------|--------------|---|
| E_1 | E'_1 | $A_1 = (E_1 \alpha_1) + (C_1 \alpha_2)$ |
| E_2 | E'_2 | $A_2 = (E_2 \alpha_1) + (C_2 \alpha_2)$ |
| E_3 | E'_3 | $A_3 = (E_3 \alpha_1) + (C_3 \alpha_2)$ |
| E_4 | E'_4 | $A_4 = (E_4 \alpha_1) + (C_4 \alpha_2)$ |
| E_5 | E'_5 | $A_5 = (E_5 \alpha_1) + (C_5 \alpha_2)$ |
| ... | ... | ... |
| E_n | E'_n | $A_n = (E_n \alpha_1) + (C_n \alpha_2)$ |

Tabela 3. Multiplicação das constantes nas previsões do AR e MLP.

A estrutura do algoritmo com a abordagem de constantes de linearidade está expressa na Figura 7.

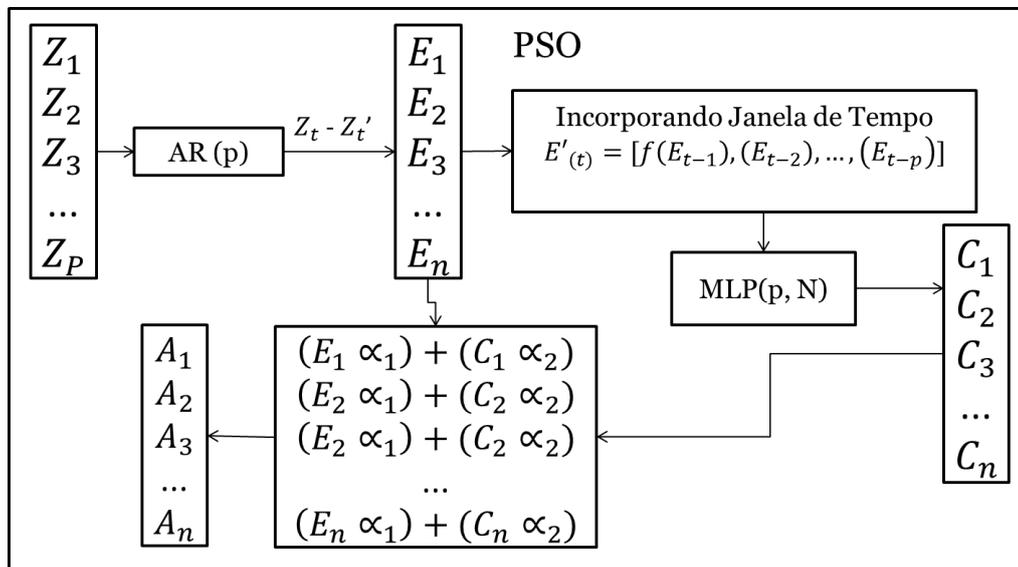


Figura 7. Estrutura do algoritmo com as constantes de linearidade.

3.3. Algoritmo de Retropropagação

Nesta abordagem presumimos que as previsões do AR e do MLP tenham uma relação não-linear e para descobrirmos essa relação utilizamos o algoritmo de Retropropagação ou BP que Segundo Barros (2011) esse nome é dado pelo fato do erro ser retropropagado durante o seu treinamento.

O BP é um algoritmo estático supervisionado por que suas entradas são conhecidas como também a saída esperada. A aprendizagem ocorre pela correção dos erros através dos ajustes de pesos. Essa aprendizagem ocorre em duas fases:

1. Fase Forward ou de propagação e;
2. Fase Backward ou de Retropropagação.

Segundo Barros (2011) na fase 1 um conjunto de sinais é submetido à camada de entrada e o seu efeito é propagado através da rede passando por todas as camadas até chegar na saída onde o resultado é fornecido.

Na fase 2 é onde o aprendizado ocorre por meio do ajuste de pesos. Para corrigir o erro obtido na fase 1 este erro é calculado pela subtração do resultado fornecido com a esperada como na Equação 3, propagando o erro para trás.

Nessa abordagem utilizamos somente uma camada escondida, a taxa de aprendizado e quantidade de neurônios foram feitas de modo empírico. A estrutura do algoritmo utilizado é ilustrado na Figura 8

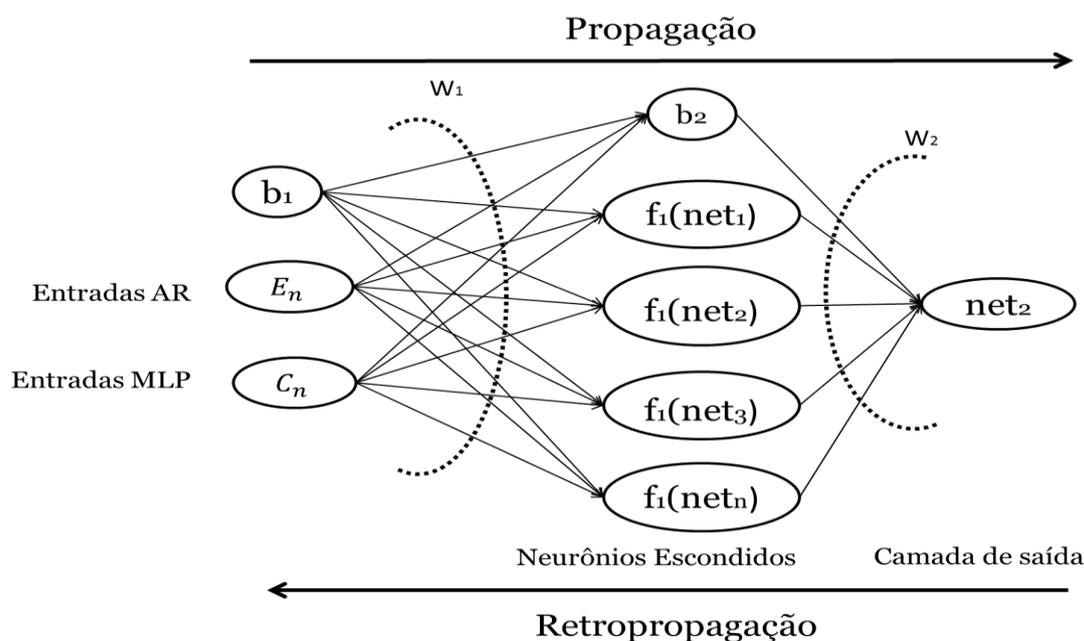


Figura 8. Estrutura do algoritmo BP.

A estrutura do BP recebe como entrada as previsões do AR e MLP. Cada elemento correspondente é colocado como entrada a fim de encontrar uma relação entre as previsões e a saída esperada. O processo de treinamento é feito para todos os elementos das previsões.

A aptidão do algoritmo PSO nesta abordagem é feita de modo diferente das demais abordagens descritas nesse trabalho. Para cada conjunto de previsões tanto do AR como MLP é calculado o MSE com a Equação 10. Depois de calculado o MSE das duas previsões, a aptidão do PSO é dada pela seguinte fórmula:

$$\text{Aptidão} = \text{MSEAR} + \text{MSEMLP} \quad (11)$$

Na aptidão do PSO o erro quadrático do AR é somado com o erro quadrático do MLP no intuito de achar o melhor conjunto de soluções para as duas abordagens. Diferente das outras abordagens que executavam toda a sua computação na aptidão do PSO, nesta o algoritmo BP só será executado ao final da otimização como mostra a Figura 9 que ilustra a execução completa do algoritmo.

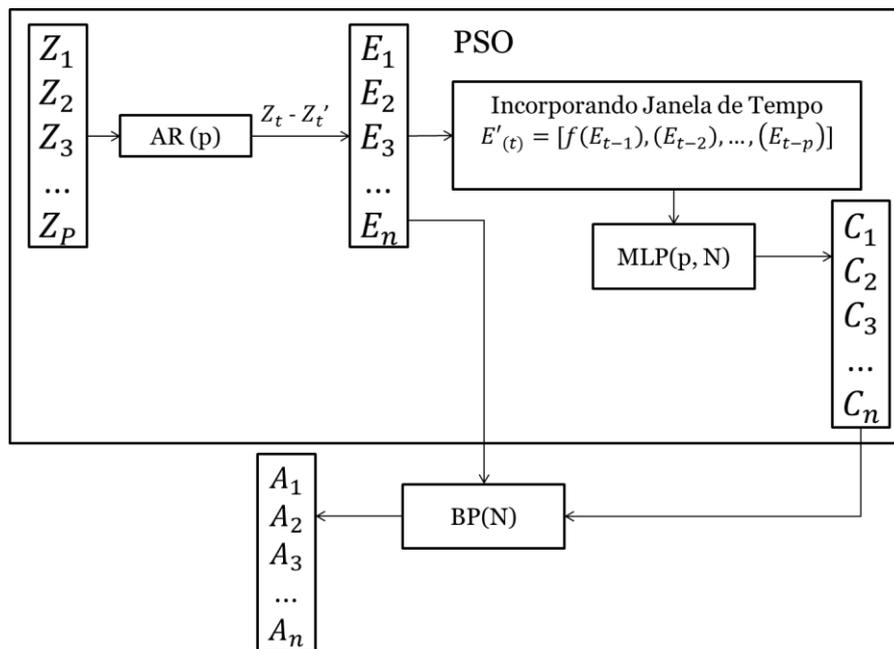


Figura 9. Estrutura do algoritmo geral com o algoritmo BP.

Para que isto ocorra são criadas duas matrizes, uma matriz de entrada que junta todas as previsões do conjunto de treinamento e teste do PSO, e na segunda matriz junta-se todas as saídas referentes ao conjunto de saída do treinamento e teste, a figura 10 ilustra essa situação. Essas duas matrizes serão a entrada e saída do algoritmo BP.

A série temporal é dividida em três partes no algoritmo BP:

1. Conjunto de treinamento, responsável em treinar os dados a fim de encontrar o melhor conjugado de pesos por meio da Retropropagação;
2. Conjunto de validação que é executado simultaneamente com o treinamento, porém nesse processo o algoritmo BP só executa a fase Forward, e;
3. Conjunto de teste que irá utilizar os pesos treinados pela rede para gerar o MSE.

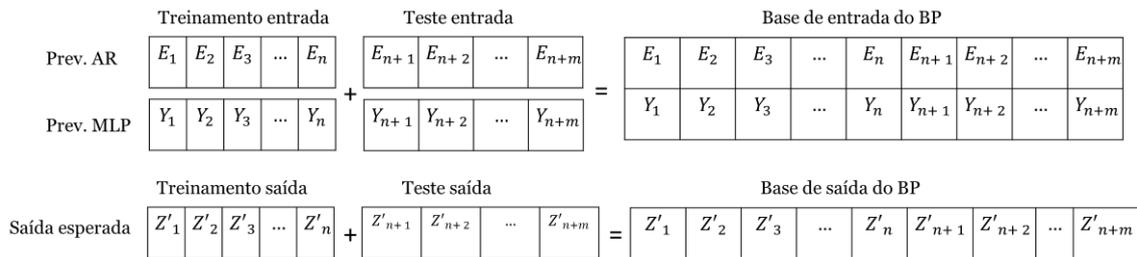


Figura 10. Base de entrada e saída do algoritmo BP.

3.4. Vantagens e Desvantagens

Técnicas de IA são amplamente utilizadas na literatura para realizar previsões de séries temporais, mas em muitos trabalhos somente um tipo de técnica de inteligência artificial é utilizado como os trabalhos de Yadav (2006) e Yi (2011). Yadav utiliza um modelo de neurônio com arquitetura polinomial que é uma variação das RNA e Yi utiliza um modelo beta autorregressivo para previsão de indicadores educacionais no Brasil.

Ambos os trabalhos conseguem realizar as previsões para os seus determinados problemas, mas Robinson (2011) afirma que a combinação de vários modelos diferentes de previsão podem melhorar a precisão das previsões em relação aos modelos individuais, além disso, o sistema que combina técnicas distintas torna-se mais robusto e eficaz em problemas que a alteração na estrutura dos dados ocorre, ou seja, se primeiramente a série temporal possuía mais elementos lineares do que não lineares um modelo de previsão linear seria eficiente nesse caso, mas se ocorrer uma alteração nos dados esse mesmo modelo não conseguirá fazer uma boa previsão como antes.

Uma das vantagens do sistema descrito nesse trabalho está relacionada a combinação de técnicas distintas de previsão. Como descrito na fundamentação teórica alguns dos adjetivos das técnicas utilizadas é que o AR possui a capacidade de modelar elementos lineares e o MLP possui a capacidade de fazer aproximações universais em dados não-lineares.

Como a composição das séries temporais é dada por elementos não lineares e lineares como segue na Equação 1. A combinação destas duas técnicas de previsão possibilita uma maior chance de modelar mais dados em relação aos

modelos individuais o tornando também robusto a alterações dos dados. Por conseguinte independente de tendência, ciclo, sazonalidade ou ruído das séries temporais a combinação destes modelos realizará bem o seu mapeamento.

Mesmo o sistema possuindo uma boa combinação de técnicas de previsão não se pode afirmar que as mesmas realizarão previsões ótimas, pois as previsões de ambos os modelos estão diretamente ligadas com os seus parâmetros, pois são estes os responsáveis em fazer com que o sistema faça o mapeamento correto das séries temporais.

O sistema pode possuir técnicas de previsão amplamente utilizadas da literatura, mas se nenhuma estratégia para a busca dos parâmetros for tomada, não podemos garantir que o sistema terá uma boa previsão. Desse modo se a escolha dos parâmetros for feita de modo tentativa erro as chances de o sistema realizar uma má previsão será alta e também consumirá muito tempo.

A utilização de um algoritmo de otimização como o PSO, que segundo Oliveira (2014) é uma possível estratégia para a definição do melhor conjunto parâmetros de técnicas de previsão. Como Prata (2012) afirma que o PSO é amplamente utilizado e aceito na literatura, devido a sua flexibilidade e robustez em problemas de otimização, isso garante que tanto o AR como o MLP irão obter ótimos pesos obtendo assim uma boa previsão para o problema trabalhado.

Portanto a otimização passa a ser outra vantagem do sistema descrito neste trabalho, pois com a combinação de todas essas técnicas a precisão das previsões tende a aumentar de forma ascendente independente do tipo de comportamento da série temporal utilizada, pois o sistema será robusto suficiente para lidar com essa problemática.

A desvantagem do sistema deste trabalho está relacionada com a complexidade computacional ao executar o algoritmo, acarretando em um maior tempo para a execução do mesmo. Mas essa desvantagem não afeta a importância e o desempenho do sistema.

4. Experimentos

Neste capítulo são apresentadas as séries temporais que foram utilizadas para os experimentos e suas características, como foram conduzidos os experimentos, os resultados e análise dos resultados.

4.1. Bases Utilizadas

Os experimentos foram realizados em 7 conjuntos de dados, 4 destes conjuntos são de séries temporais educacionais, 3 dessas bases são do United States Census Bureau, que são: Pessoas que completaram o ensino fundamental em 4 anos ou menos; Pessoas que completaram o ensino fundamental entre 5 a 8 anos e; Porcentagem de pessoas de 25 anos ou mais que concluíram a graduação. A outra base vem do Department of Education and Training – Australian Government que é a quantidade de estudantes que conseguiram concluir o ensino superior.

Para afirmar o desempenho do sistema híbrido descrito neste trabalho foram escolhidas 3 bases da biblioteca Time Series Data Library, amplamente utilizada na literatura. As bases escolhidas foram: Quantidade de passageiros de companhias aéreas; Número de manchas solares e; Quantidade de veículos e peças produzidos no Canadá.

4.1.1. Séries educacionais do United States Census Bureau.

Os dados de todas as séries educacionais advindas do United States Census Bureau foram coletadas anualmente de 1940 até 2014, possuindo um tamanho de 59 pontos devido a alguns anos não terem sido coletados. A Figura 11 representa respectivamente a quantidade de pessoas latino-americanas que completaram o ensino fundamental em 4 anos ou menos. A Figura 12 representa a quantidade de pessoas latino-americanas que completaram o ensino fundamental entre 5 a 8 anos. A Figura 13 demonstra a porcentagem de pessoas latino-americanas de 25 anos ou mais que concluíram a graduação.

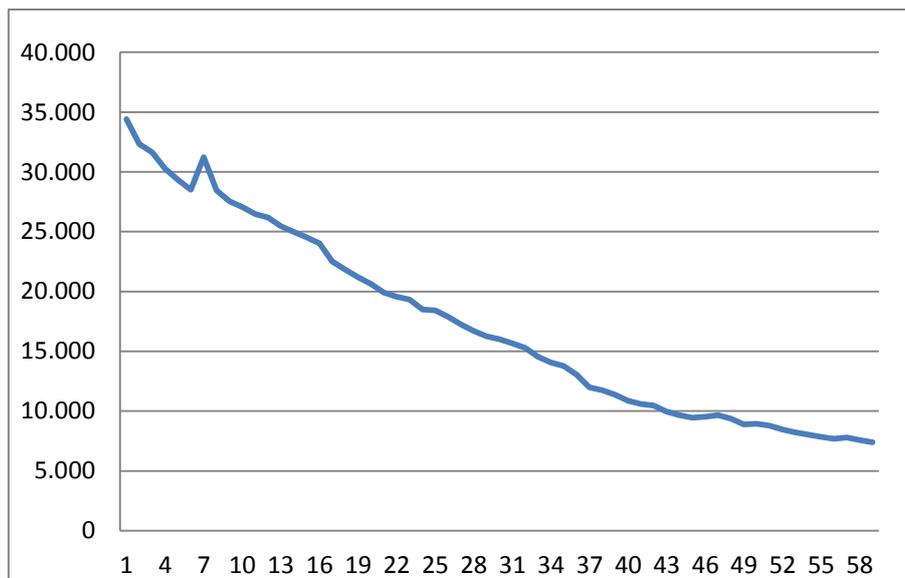


Figura 11. Pessoas latino-americanas que completaram o ensino fundamental em 4 anos ou menos.

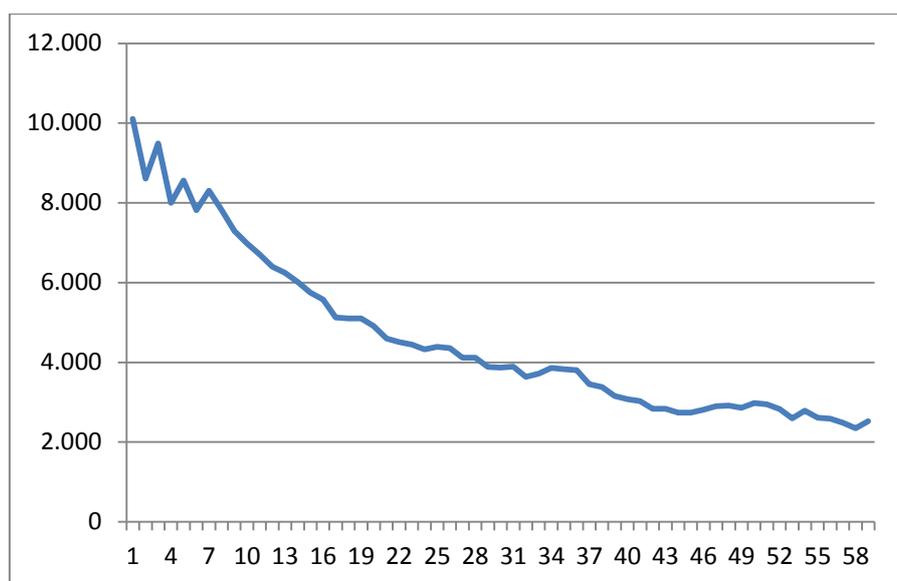


Figura 12. Pessoas latino-americanas que completaram o ensino fundamental entre 5 a 8 anos.

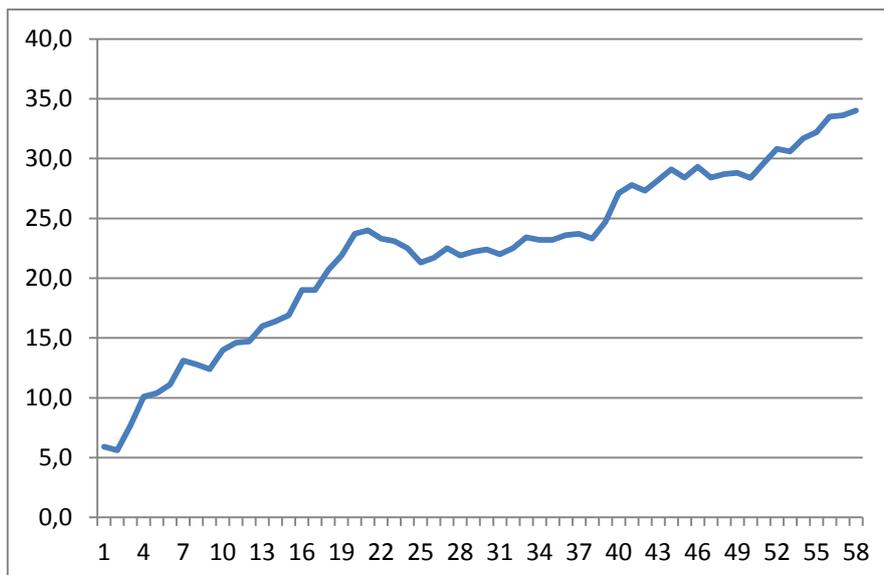


Figura 13. Porcentagem de pessoas latino-americanas de 25 anos ou mais que concluíram a graduação.

4.1.2. Quantidade de estudantes que conseguiram concluir o ensino superior.

Esta série educacional apresenta estatísticas sobre estudantes de instituições de ensino superior da Austrália. Foram coletadas anualmente a quantidade de estudantes que conseguiram concluir o ensino superior a partir de 1949 até 2000, possuindo um tamanho de 51 pontos como demonstra a Figura 14.

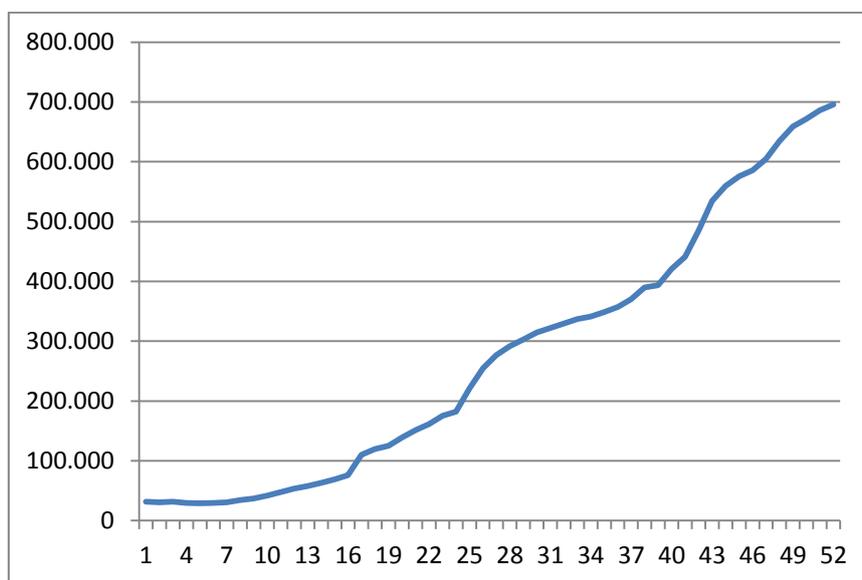


Figura 14. Quantidade de concluintes de ensino superior na Austrália.

4.1.3. Quantidade de passageiros de companhias aéreas.

Esta série temporal é referente a quantidade de passageiros das companhias aéreas internacionais. As observações foram coletadas mensalmente desde janeiro de 1949 a dezembro de 1960, possuindo 144 pontos como demonstra a Figura 15.

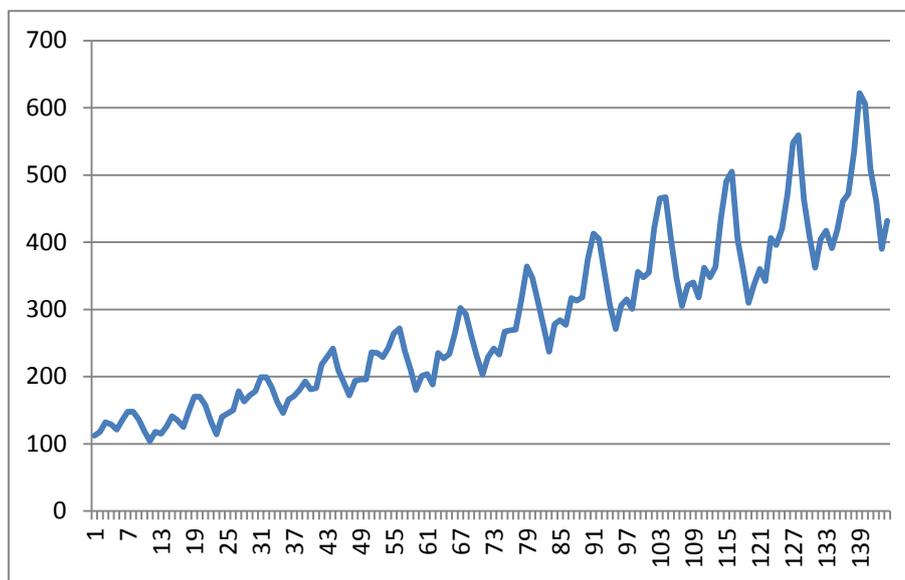


Figura 15. Quantidade de passageiros das companhias aéreas internacionais.

4.1.4. Número de manchas solares.

Os dados das manchas solares considera um número anual de 1700 a 1987 contendo um total de 288 pontos. Segundo Zhang (2001) o estudo da importância das manchas solares tem grande importância para geofísicos, cientistas e climatologistas. O enredo desta série temporal sugere que existe um ciclo médio de cerca de 11 anos. A Figura 16 ilustra essa série temporal.

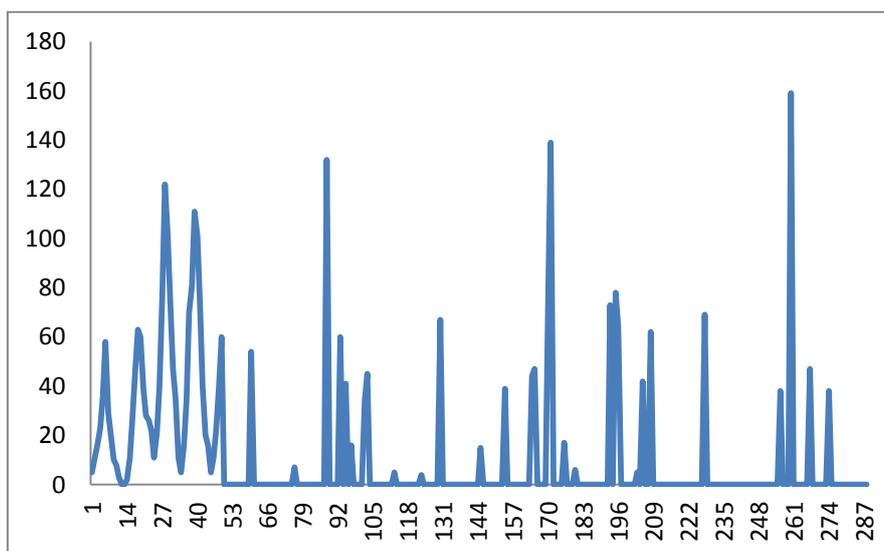


Figura 16. Número de manchas solares entre 1700 a 1987.

4.1.5. Quantidade de veículos e peças produzidos no Canadá.

Esta série está relacionada com a quantidade de vendas de veículos, motores e peças no Canadá. Os dados desta série temporal foram coletadas mensalmente no Canadá desde janeiro de 1976 a agosto de 1991, contendo um tamanho de 188 pontos, como demonstra na Figura 17.

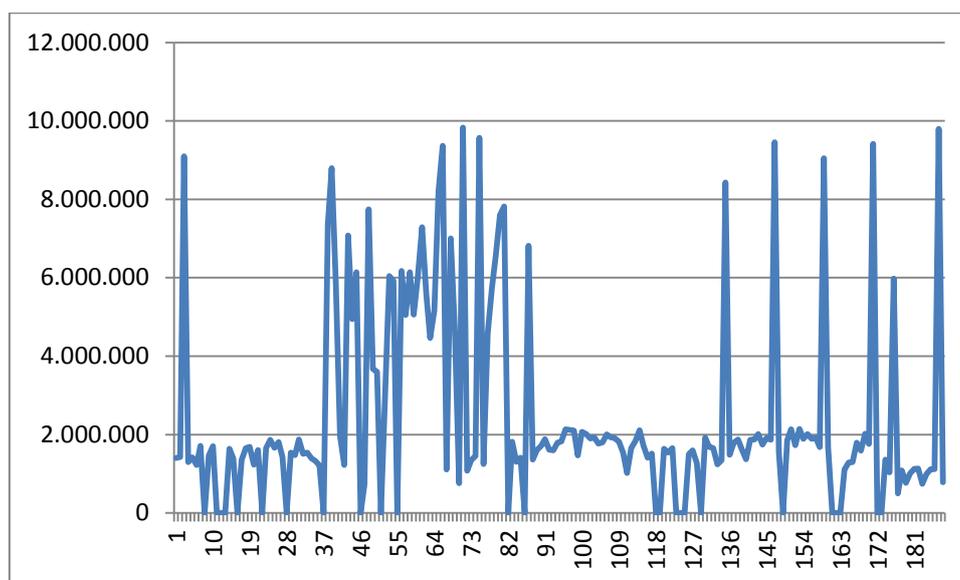


Figura 17. Quantidade de veículos e peças vendidos no Canadá.

4.1.6. Tabelas com abreviações das séries temporais.

Para melhor entendimento do processo de experimentos e análise dos resultados, a Tabela 4 resume algumas características dos conjuntos de dados utilizados:

| Nome | Descrição | Pontos |
|------------------------------|--|-----------|
| Séries Educacionais | | |
| Conclusão 4 anos | Quantidade de pessoas latino-americanas que completaram o ensino fundamental em 4 anos ou menos. Dados coletados entre 1940 e 2014 | 59 |
| Conclusão 5~8 anos | Quantidade de pessoas latino-americanas que completaram o ensino fundamental entre 5 e 8. Dados coletados entre 1940 e 2014 | 59 |
| Concluintes Graduação | Porcentagem de pessoas latino-americanas de 25 anos ou mais que concluíram a graduação. Dados coletados entre 1940 e 2014 | 59 |

| | | |
|---------------------------------|--|------------|
| Concluintes Austrália | Quantidade de concluintes de ensino superior na Austrália. Dados coletados entre 1949 e 2000. | 51 |
| Séries de Áreas Diversas | | |
| Aéreas | Quantidade de passageiros das companhias aéreas internacionais. Dados coletados mensalmente entre janeiro de 1949 a dezembro de 1960. | 144 |
| Manchas | Quantidade de manchas solares coletadas mensalmente entre o 1700 a 1987. | 288 |
| Veículos | Quantidade de vendas de veículos, motores e peças no Canadá. Dados coletados mensalmente no Canadá desde janeiro de 1976 a agosto de 1991. | 188 |

Tabela 4. Resumo das características das séries temporais utilizadas.

4.2. Configuração de experimentos

O sistema híbrido deste trabalho é comparado com outros métodos, tais como o MLP tradicional utilizando o algoritmo BackPropagation, PSO-AR, PSO-MLP, PSO-AR-MLP, PSO-AR-MLP-Linear e PSO-AR-MLP-BP. Para todas as abordagens a implementação foi realizada na linguagem de programação Java. Os conjuntos de dados foram repartidos usando 60% para treinamento, 20% para validação e 20% para teste. A Figura 18 demonstra como foi particionado a série temporal.

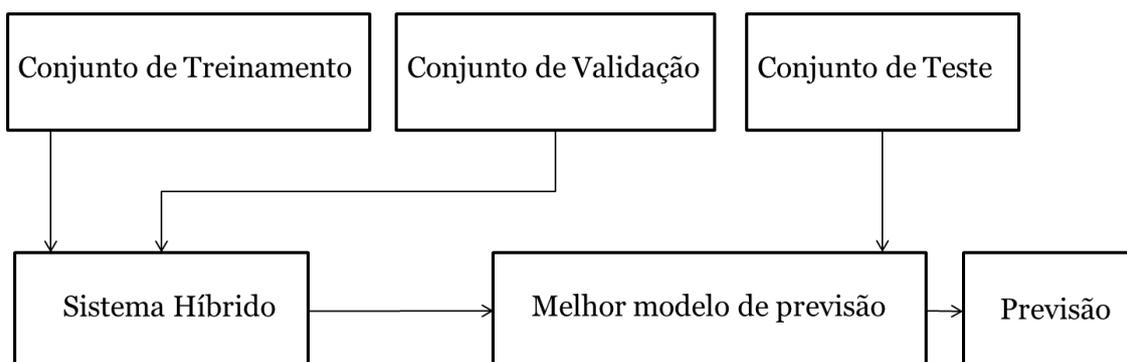


Figura 18. Fluxo do experimento.

No método MLP com o algoritmo BackPropagation o algoritmo foi executado por 1000 iterações, utilizando para a execução do experimento 10 neurônios na sua camada escondida, e como taxa de aprendizado uma constante com valor igual a 0,02.

Para os métodos que fazem uso do PSO, o algoritmo é responsável em realizar uma busca estocástica buscando o melhor conjunto de parâmetros para cada técnica de previsão, e para isto foi definido um valor padrão. O algoritmo é executado por 1000 iterações, tendo o seu enxame com 30 partículas, peso de inércia 0.5 e usando seus fatores de aprendizagem com $c_1 = 2$, e $c_2 = 2$.

Os métodos que utilizam o modelo de previsão AR a ordem do modelo foi definida com 5 observações passadas e para os métodos que utilizam a arquitetura do MLP foi usado 10 neurônios para sua camada escondida e taxa de aprendizado igual a 0,02.

Para todos os conjuntos de dados utilizados no experimento, inicialmente realizou-se a normalização dos seus elementos no intervalo de (0 e 1), com o objetivo de coloca-los em um domínio pré-determinado. Para esta normalização a seguinte equação foi utilizada:

$$\tilde{X}_n = \frac{X_0 - X_{min}}{X_{max} - X_{min}}, \quad (12)$$

Onde X_n representa o valor resultante da série, X_0 representa o primeiro valor da série, X_{min} o valor mínimo e X_{max} o valor máximo.

Para todas as abordagens os experimentos foram executados por 30 iterações, a média MSE foi realizada e pode ser analisada na Figura 20.

4.3. Análise de resultados

As previsões foram realizadas um passo a frente, buscando saber qual o próximo valor que possivelmente seguirá a série temporal. Os experimentos foram aplicados a 7 séries temporais, sendo 4 séries educacionais e 3 séries de áreas diversas, utilizando 5 métodos distintos de previsão, com o objetivo de verificar o modelo que apresentasse o melhor desempenho.

A Figura 20 mostra a média do erro em relação a quantidade de execuções de cada técnica, o erro é calculado a partir do conjunto de testes produzido. Os valores de erros mais baixos estão em destaque na Tabela 5.

| Método | Conclusão 0 4 anos | Conclusão 5~8 anos | Concluintes Graduação | Concluintes Austrália | Aéreas | Manchas | Veículos |
|---------|--------------------|--------------------|-----------------------|-----------------------|---------|---------|----------|
| MLP | 0.0026 | 0.0027 | 0.01049 | 0.2718 | 0.0667 | 0.04743 | 0.0399 |
| PSO_AR | 6.2959E-4 | 5.7192E-4 | 3.5651E-5 | 0.0024 | 0.01166 | 0.0083 | 0.02707 |
| PSO_MLP | 0.0059 | 0.0016 | 0.0025 | 0.0728 | 0.0228 | 0.01167 | 0.02523 |

| | | | | | | | |
|--------------------------|------------------|------------------|-----------|------------------|------------------|------------------|------------------|
| PSO_AR_MLP | 0.0041 | 0.0042 | 0.0024 | 0.0012 | 5.4254E-4 | 1.4840E-4 | 3.3370E-4 |
| PSO_AR_MLP_LINEAR | 1.7204E-4 | 6.8768E-5 | 3.8141E-4 | 4.9183E-4 | 1.5488E-4 | 9.0574E-4 | 3.6045E-4 |
| PSO_AR_MLP_BP | 0.05962 | 0.0881 | 0.08654 | 0.0469 | 0.0941 | 0.0224 | 0.3060 |

Tabela 5. Tabela de Experimentos.

De acordo com os experimentos gerados 3 das 5 técnicas se destacaram, o PSO_AR obteve melhor desempenho na base Concluintes Graduação, o PSO_AR_MLP também obteve melhor desempenho na base Veículos e o PSO_AR_MLP_LINEAR obteve melhor desempenho em 5 dos conjuntos de dados (Conclusão 4 anos, Conclusão 5~8 anos, Concluintes Austrália, Aéreas, Manchas).

Ao analisar os erros obtidos pode-se observar que as técnicas não-lineares como o MLP, PSO_MLP, PSO_AR_MLP_BP conseguiram mapear os padrões dos dados, porém não obtiveram resultados tão satisfatórios quanto aos métodos lineares como o PSO_AR, PSO_AR_MLP e PSO_AR_MLP_LINEAR. Por conseguinte para as bases de dados utilizadas neste trabalho as técnicas lineares conseguiram realizar previsões mais precisas.

Mesmo o modelo sugerido neste trabalho (PSO_AR_MLP_LINEAR) não obtendo as melhores previsões em todas as séries a análise do erro mostra que as previsões chegam tão próximas quanto as técnicas que obtiveram os melhores resultados, demonstrando que a técnica proposta consegue mapear todos os padrões em todos os conjuntos de dados afirmando a sua robustez sobre qualquer tipo de dados.

As seguintes imagens ilustram o conjunto de treinamento e teste das previsões geradas pelo PSO_AR_MLP_LINEAR. As previsões são comparadas com as séries originais:

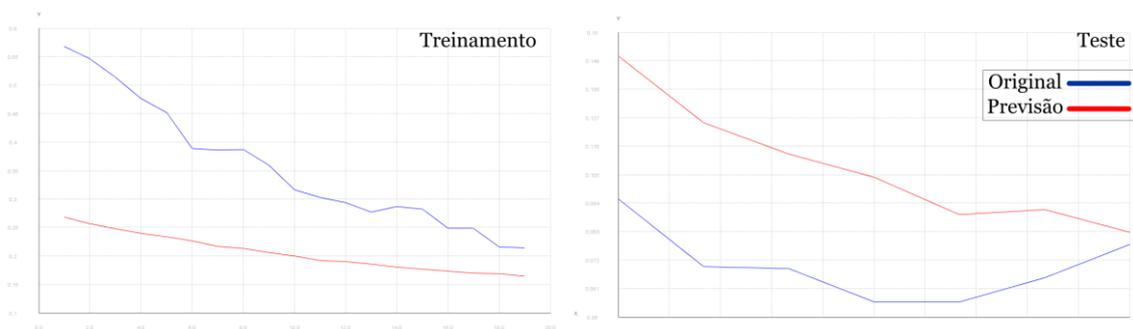


Figura 19. Previsão do modelo proposto em comparação com a série Conclusão 4 anos.

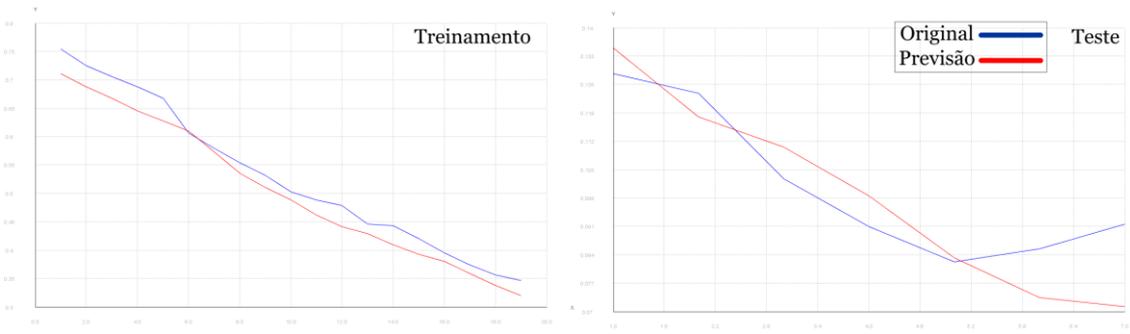


Figura 20. Previsão do modelo proposto em comparação com a série Conclusão 5 a 8 anos.

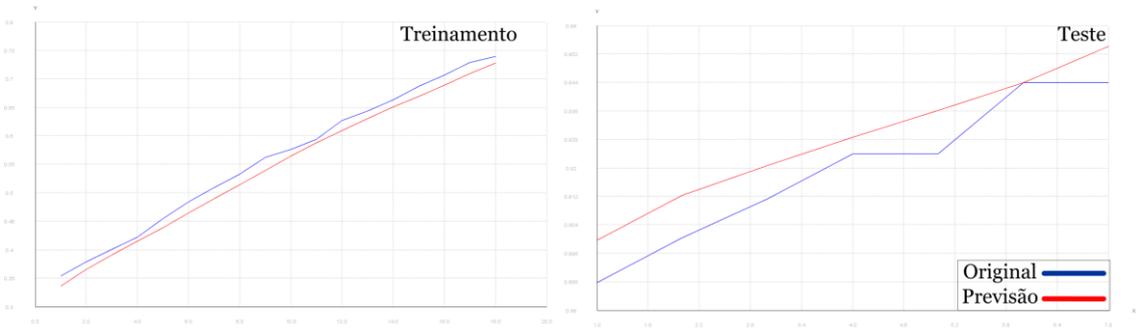


Figura 21. Previsão do modelo proposto em comparação com a série Conclusão Graduação.

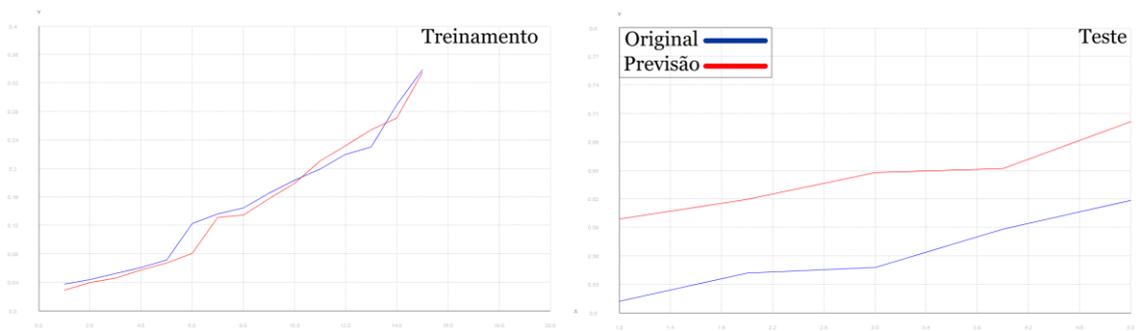


Figura 22. Previsão do modelo proposto em comparação com a série Conclusão Austrália.

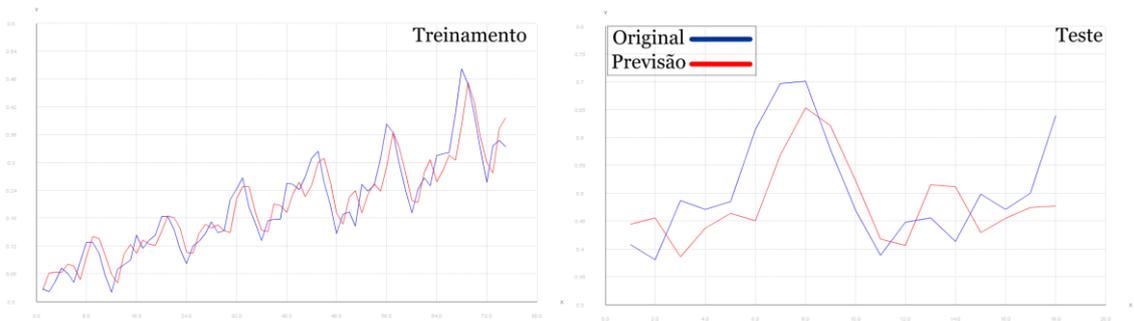


Figura 23. Previsão do modelo proposto em comparação com a série Aéreas.

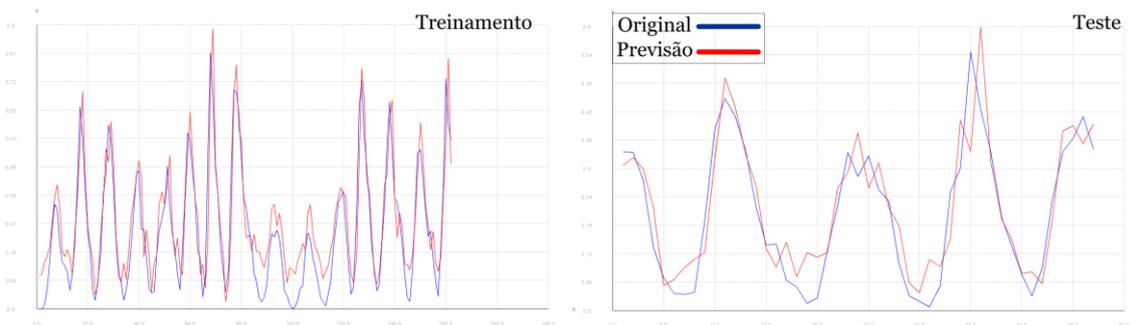


Figura 24. Previsão do modelo proposto em comparação com a série Manchas.

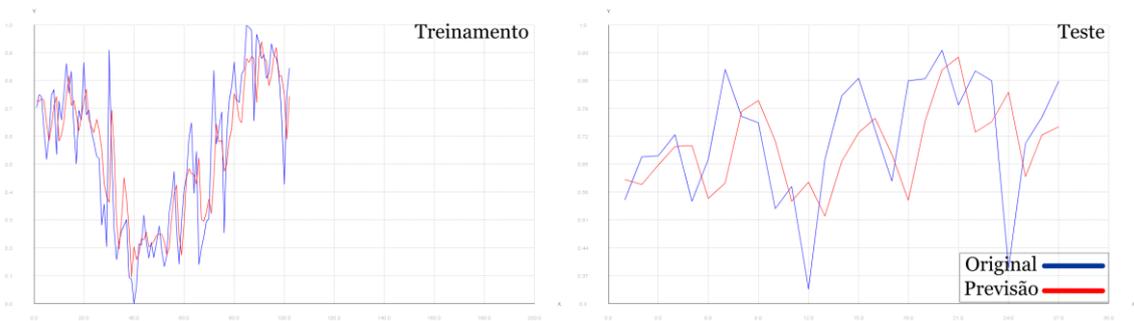


Figura 25. Previsão do modelo proposto em comparação com a série Veículos.

Portanto, com a análise dos dados podemos afirmar que a combinação de modelos que possuem forças distintas de previsões conseguem ser mais eficientes comparado a modelos únicos e simples.

5. Conclusões e Trabalhos Futuros

Neste capítulo são apresentadas as considerações finais e a sugestão de trabalhos futuros.

5.1. Conclusões

Este trabalho teve por objetivo o desenvolvimento de um sistema híbrido capaz de gerar previsões em séries temporais educacionais. Espera-se que essas previsões desperte interesse aos gestores da educação, para que dessa forma eles possam traçar metas, planos e estratégias para a criação de ações de combate as taxas futuras.

O sistema é composto por um modelo autorregressivo para a previsão de dados lineares, um modelo de previsão não-linear a técnica MLP e um algoritmo de otimização por enxame de partículas, para encontrar o melhor conjunto de parâmetros para ambos os métodos.

O sistema teve como vantagens a combinação de modelos distintos e a otimização em busca de melhores parâmetros, aumentando a capacidade de previsão do sistema, bem como a sua robustez a mudança de dados.

Para treinar e testar a eficiência do modelo proposto um conjunto de 7 séries temporais foram utilizadas, sendo 4 bases de séries temporais educacionais e 3 séries de áreas diversas. As séries temporais educacionais compreendem a quantidade de pessoas que completaram o ensino fundamental em 4 anos ou menos, quantidade de pessoas que completaram o ensino fundamental entre 5 a 8 anos, porcentagem de pessoas de 25 anos ou mais que concluíram a graduação e a quantidade de estudantes que conseguiram concluir o ensino superior.

As séries temporais de áreas diversas utilizadas para afirmar o modelo compreendem a quantidade de passageiros de companhias aéreas internacionais, número de manchas solares e quantidade de veículos e peças produzidos no Canadá.

Para análise do desempenho da técnica foi utilizado um conjunto de 5 técnicas distintas de previsão a fim de afirmar a sua eficiência em relação aos demais modelos. O método proposto e a sua variação obteve os melhores resultados em 6 dos 7 conjuntos de dados utilizados para a experimentação, demonstrando resultados promissores. A análise dos dados mostrou que para o conjunto de dados utilizados neste trabalho as técnicas lineares obtiveram melhores resultados.

5.2. Trabalhos Futuros

O sistema híbrido utilizado neste trabalho é capaz de realizar previsões em qualquer série temporal estática, na qual os dados utilizados no modelo não se alteram. Os resultados dessa pesquisa são bastantes satisfatórios, porém para tornar esta metodologia eficiente em diversos tipos de ambientes de dados sugerimos a implementação de adaptações para ambientes dinâmicos que possuem mudança de conceito.

Vislumbra-se o desenvolvimento de um modelo de previsão que seja capaz de lidar com o fluxo de dados dinâmicos, onde os dados chegam de pedaço em pedaço e também com mudança de conceito (Concept Drift), onde os dados se alteram ao longo do tempo (CAVALCANTE, 2015).

Referências

ADHIKARI, R. "A Neural Network Based Linear Ensemble Framework for Time Series Forecasting" *Neurocomputing*, 2015.

ANDRADE. N. S e CARVALHO. P. S. "Análise das Variações de Temperaturas na Plataforma de Nossa Senhora da Glória Via Box E Jenkins" *GEINTEC. Gestão, Inovação e Tecnologias*. 2011, p. 54-48.

BARROS. A. C. A. "Otimização de Redes Neurais Para Previsão de Séries Temporais" 2011.

Censo Escolar da Educação Básica 2012 Resumo Técnico. Disponível em: http://download.inep.gov.br/educacao_basica/censo_escolar/resumos_tecnico_s/resumo_tecnico_censo_educacao_basica_2012.pdf Acesso em: 08 set. 2015.

Cavalcante. C. R., Oliveira. A. L. I. "An Approach to Handle Concept Drift in Financial Time Series Based on Extreme Learning Machines and Explicit Drift Detection. *IEE*, 2015.

Department of Education and Training – Australian Government. Disponível em <<http://docs.education.gov.au/node/35519>>

FERNANDES, L. G. L, PORTUGAL, S. M. e NAVAU, P. O. A "Previsão de Séries de Tempo: Redes Neurais Artificiais e Modelos"

LATORRE. M. R. D. O e CARDOSO. M. R. A. "Análise de séries temporais em epidemiologia: uma introdução sobre os aspectos metodológicos" Disponível em: <<http://www.scielo.br/pdf/rbepid/v4n3/02.pdf>> Acesso em: 09 set. 2015.

LEE. Y S. e TONG. L.I. "Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming" *Elsevier*, 2010.

MAIA, A. L. S e NETO, F. C., "Dinâmica inflacionária brasileira: resultados de auto-regressão" 2006.

OLIVEIRA, J. F. L. e LUDERMIR, T. B. A Hybrid Evolutionary System for Parameter Optimization and Lag Selection in Time Series Forecasting. In: 2014 Brazilian Conference on Intelligent Systems (BRACIS), 2014, Sao Paulo. 2014 Brazilian Conference on Intelligent Systems. p. 73.

PORTAL INEP. "O que é Ideb?" Disponível em: <<http://portal.inep.gov.br/web/portal-ideb/o-que-e-o-ideb>> Acesso em: 08 set. 2015.

PRATA, B. A. "Um algoritmo enxame de partículas para uma variante do problema de máxima cobertura". *GEPROS. Gestão da Produção, Operações e Sistemas*, Ano 7, nº 2, abr-jun/2012, p. 139-148.

- R. Hyndman. (2010) Time series data library. [Online]. Disponível em: <<http://data.is/TSDLdemo>>
- ROBINSON. J. M. "Previsão de séries temporais usando ARIMA híbrido e modelo de rede neural" 2001.
- RIBEIRO. G. H. T, NETO. P. S. G. M, CAVALCANTI. G. D. C. e TSANG. I. R. "Lag Selection for Time Series Forecasting using Particle Swarm Optimization" IEE. 2011.
- SILVA, M. I, GUIMARÃES. E. C e TAVARES. M. "analisar a série de temperatura média mensal da cidade de Uberlândia, MG, descrevendo seus componentes, e fazer previsões para períodos subsequentes através de modelos ajustados para a série." 2005.
- TSENG. F.M, YU. H.C. E TZENG. G.H. "Combining neural network model with seasonal time series ARIMA model" Technological forecasting and social change, 2000.
- United States Census Bureau. Disponível em: <<http://www.census.gov/hhes/socdemo/education/data/cps/historical/>>
- WERNER, L. e RIBEIRO, J. L. D. "previsão de demanda: uma aplicação dos modelos boxjenkins na área de assistência técnica de computadores pessoais" 2006.
- YOLCU. U, Egrioglu. E. e ALADAG. C. H. "A new linear & nonlinear artificial neural network model for time series forecasting" Elsevier, 2012.
- YI. T. L. "Aplicação de um modelo Beta Auto-Regressivo a indicadores educacionais do Brasil" Juiz de Fora. 2011.
- YADAV. R. N, KALRA. P.K e JOHN. J. "Times series prediction with single multiplicative neuron model" ScienceDirect. 2006.
- ZHANG. G. P. "Time series forecasting using a hybrid ARIMA and neural network model" Neurocomputing, 50. 2003. pp. 159-175.