

# Uma aplicação de autoencoders para a camada física

RESUMO EXPANDIDO - Disciplina de TCC290009

**Lucas Thiesen**

Estudante do Curso de Engenharia de Telecomunicações

**Roberto Wanderley da Nobrega**

Professor orientador

Semestre 2019-1

***Resumo-** Cada vez mais aumenta a demanda por eficiência do uso de canais sem fio, e com o decorrer do tempo, várias técnicas surgiram para otimizar a utilização desse meio de transmissão. Uma variedade de códigos corretores de erro, diferentes técnicas de modulação, inúmeras formas de uso da banda do canal e uma gama de métodos de acesso ao meio.*

*Na indústria está se vendo diversas aplicações de aprendizado de máquina: Processamento de imagens, previsão de comportamento, marketing, etc. Diferente dessas áreas se vê muito poucas aplicações para sistemas de telecomunicação. Recentemente uma nova linha de pesquisa surgiu, modelar o sistema de comunicação ponta a ponta como um autoencoder, e vários resultados interessantes foram alcançados para canais teóricos.*

*Nesse trabalho será apresentado uma linha de raciocínio levemente diferente. Ao invés de treinar uma rede neural com base em um canal teórico, será feito o treinamento do modelo sobre um canal real. Esse sistema deve aprender a lidar com os diferentes contrastes de um canal sem fio, e dessa forma atingir uma alta eficiência no uso do canal de comunicação. Ao final do trabalho, os resultados do sistema treinado utilizando o ambiente descrito previamente, será comparado as técnicas que vem sendo empregadas popularmente na indústria e literatura.*

**Palavras-chave:** deep learning. Wireless communication. Software defined radio (SDR).

## 1 Introdução

Todos os dias surgem novos cenários, aplicações e desafios para redes de comunicação sem fio. É esperado que a próxima geração de redes sem fio suporte taxas de transmissão extremamente altas e aplicações radicalmente novas, e isso irá forçar com que novos

paradigmas de comunicação sem fio surjam (JIANG et al., 2017). É esperado que em 2020, 50 bilhões de dispositivos estejam conectados com a nuvem (GUPTA; JHA, 2015), isso traz novos desafios técnicos. Alguns deles são de taxas de transmissão média de 10 a 100 vezes maiores, latência ponta a ponta reduzida em até 5 vezes, volume de dados por área 1000 vezes maior (OSSEIRAN et al., 2014; GUPTA; JHA, 2015). Todas essas inferências só ressaltam, é necessário procurar novos meios de fazer comunicação sem fio, meios mais eficientes.

A camada física da rede, sendo um campo muito maduro e especializado nos estudos de telecomunicações, vem tendo aumentos de desempenho cada vez menores, cada vez mais esses sistemas se aproximam do limite. Diversas camadas da rede já viram benefícios no uso de técnicas de *machine learning* e *deep learning* para ter melhoras de desempenho, porém, a camada física vem se aproveitando muito pouco desse novo campo de estudo (O'SHEA; HOYDIS, 2017).

Recentemente com maior desenvolvimento ao redor de padrões como o 5G e com uma demanda crescente de sistemas de comunicação sem fio mais performáticos, começa a surgir com maior volume pesquisas e aplicações de *deep learning* e *machine learning* na camada física. Seja para tornar mais eficiente e dinâmica a rede, seja para melhorar o desempenho e taxas de transmissão ou até para melhorar a adaptabilidade da rede ao canal. E a expectativa é que em breve a camada física começará a colher os frutos dessa tecnologia emergente.

As técnicas clássicas de comunicação sem fio separam os sistemas em vários blocos, todos altamente otimizados dentro seu próprio domínio. Mas não é claro que um conjunto de blocos, onde cada bloco é ótimo, fará com que o sistema em sua totalidade se torne ótimo. Além disso, embora muitos estudos recentes usem técnicas de *machine learning* ou *deep learning* para melhorar o desempenho dos sistemas de comunicação, poucas pesquisas foram feitas quanto a tornar o sistema de comunicação um bloco único a ser otimizado (O'SHEA; HOYDIS, 2017).

Outro ponto é que a maioria dos algoritmos utilizados em sistemas de comunicação hoje, são comprovadamente ótimos. Entretanto, a maioria desses modelos são lineares, estacionários e presumem um canal Gaussiano. O grande problema é que canais reais, normalmente, não seguem essas características completamente (O'SHEA; HOYDIS, 2017; SCHENK, 2008). Isso junto à todas as imperfeições de hardware formam um conjunto de incertezas quanto ao modelo utilizado (O'SHEA; HOYDIS, 2017; RAPPAPORT, 2002). O que é outro argumento a favor de um sistema modelado como um bloco único, que é otimizado ponta a ponta.

Inspirados pelo sucesso que alguns modelos de aprendizado de máquina possuem em processamento de imagens, modelos como redes neurais convolucionais (CNN) e autoencoders. Foi realizado o treinamento de sistemas de redes neurais que aprendem a se comunicar por um canal Gaussiano e possuem resultados muito satisfatórios e comparáveis as técnicas largamente utilizadas hoje (O'SHEA; HOYDIS, 2017). A característica mais interessante de um sistema como esse é, que diferente dos métodos convencionais, uma rede neural consegue se adaptar e aprender a se comunicar em qualquer canal que seja utilizado no treinamento, mesmo que não seja conhecido nenhum esquema teoricamente

ótimo para o mesmo (O'SHEA; HOYDIS, 2017).

Baseado nos trabalhos de O'Shea e Hoydis (2017), o foco deste trabalho é criar um modelo dinâmico de um sistema de comunicação ponto a ponto, estabelecido sobre um canal sem fio real. Este sistema deve ser capaz de aprender a se comunicar e a lidar com as variações do canal se baseando em diferentes técnicas de aprendizado de máquina. Ao final o desempenho do sistema será comparado ao das metodologias empregadas amplamente na literatura, para determinar a eficácia do modelo.

## 2 Metodologia

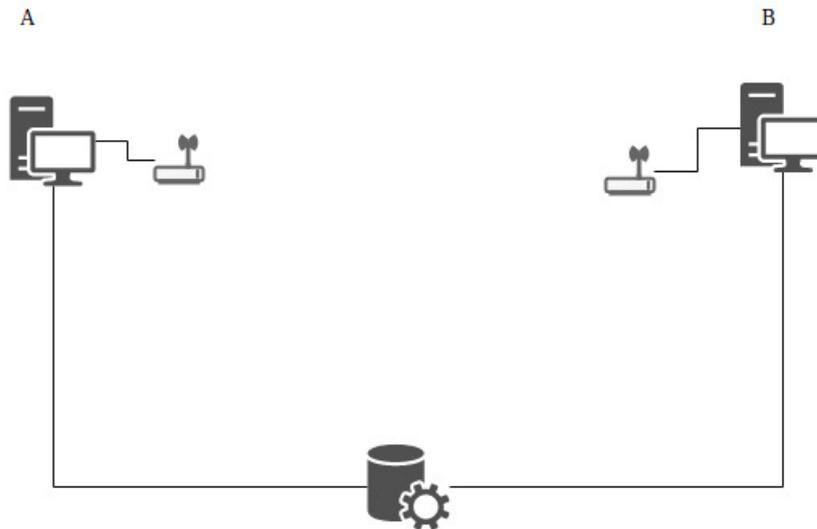


Figura 1 - Diagrama da proposta base

O objetivo inicial do projeto é criar uma aplicação que seja capaz de usar algoritmos de aprendizado de máquina para a criação de um modelo de comunicação entre os sistemas A e B.

A aplicação constitui de dois computadores cada um conectado a um dispositivo SDR e antena, e um servidor encarregado de executar o algoritmo de treino.

### 2.1 *Autoencoders*

*Autoencoders* são um subconjunto de uma classe de redes neurais conhecida como *deep belief network*, tais redes neurais são modelos generativos probabilísticos compostos de  $n$  camadas de variáveis estocásticas. Essas redes possuem uma forma muito eficiente de aprendizado camada a camada, onde a camada posterior usa o valor das variáveis da camada anterior para o processo de treinamento, e isso se propaga pela rede de forma que a primeira camada é treinada a partir dos próprios dados. Esse eficiente modelo de treinamento combinado com outras técnicas de treino e otimização possuem a capacidade de melhorar significativamente o desempenho do modelo gerado pela rede neural (HINTON, 2009).

Um *autoencoder* é uma arquitetura muito singular, onde os dados de entrada da rede são codificados e comprimidos até uma representação desejada para então passar por um processo de decodificação onde o *autoencoder* tenta inferir a partir da forma codificada

qual era seu valor original. Tal estrutura é frequentemente usada para reconhecimento de padrões em imagens, estruturas textuais e transformações áudio texto e vice-versa (BENGIO et al., 2009).

Devido a essa arquitetura singular, o *autoencoder* se torna uma estrutura muito atrativa para resolver o problema proposto por este artigo. Onde a única diferença do modelo descrito é que haverá uma camada responsável pela distorção do sinal entre a estrutura codificadora e a estrutura decodificadora, o canal.

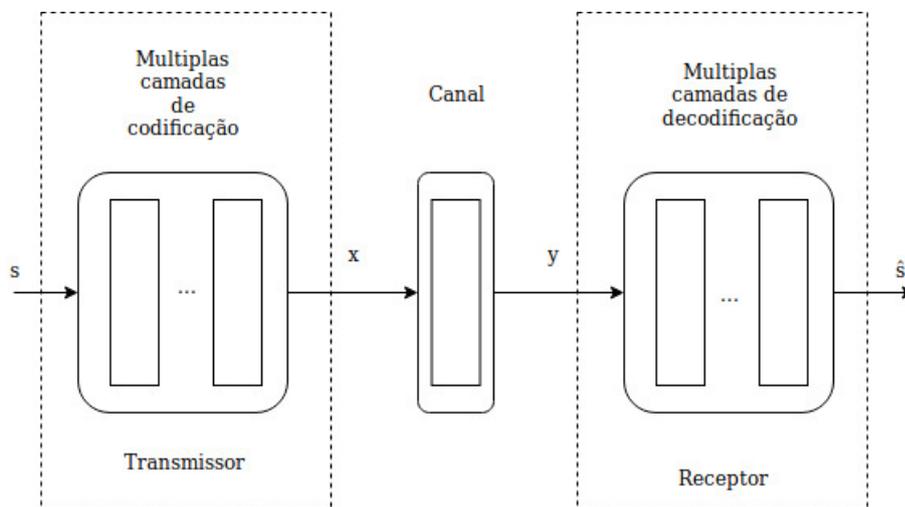


Figura 2 - Arquitetura de *autoencoder* proposta

## 2.2 Implementação inicial

Para a comunicação sem fio será usado tecnologias de rádio definido por software como o GNU Radio. O GNU Radio permite a criação de blocos de software que efetuam o processamento de sinais necessário para a transmissão da informação via rádio, e cuida de todo o processo de comunicação com o dispositivo SDR.

Para realizar o treino do sistema será construído um algoritmo de treino com base em autoencoders. Porém o autoencoder precisará ser separado em três partes.

- As camadas de codificação do autoencoder ficarão separadas das camadas decodificadoras e sua saída será transmitido para o subsistema A;
- A camada de ruído do autoencoder será substituída pelo próprio canal de comunicação entre os pontos A e B;
- As camadas de decodificação do autoencoder ficarão separadas das camadas codificadoras e receberão como entrada os dados tratados no subsistema B;

A ideia inicial para o subsistema A é criar um bloco GNU Radio que não possui inteligência, sua única função é pegar as informações da saída da parte codificadora do autoencoder e fazer o tratamento necessário para que o dado seja transmitido na rede sem fio.

E no subsistema B seria necessário criar um bloco que fizesse o processo contrário. Receber a informação pelo rádio, tratar a informação e enviá-la de volta ao servidor como

entrada para a parte decodificadora do autoencoder, finalizando ciclo de treinamento da rede neural.

## 2.3 Treino

Para o conjunto de treino será repetido um grande numero de combinações de símbolos diversas vezes, por um longo período de tempo. A ideia é que o sistema além de aprender a se comunicar, aprenda a lidar com variações no canal, diferentes níveis de atenuação e ruído e qualquer outra dificuldade que surgir no período de treino. Isso irá reforçar a capacidade do sistema de se adaptar ao canal no qual ele foi inserido.

## 2.4 Avaliação de desempenho

Será feita uma avaliação de desempenho dos modelos criados por diferentes estruturas de autoencoders e será comparado a técnicas de modulação, codificação e equalização, utilizadas em larga escala atualmente.

Os parâmetros avaliados serão, taxas de transferência alcançadas, taxa de erro de bit e taxa de erro de bloco.

## 2.5 Possíveis otimizações e refinamentos na implementação original

Em questão de aperfeiçoamento da implementação inicial, há vários caminhos a serem seguidos. São alguns deles: aprimoramento do modelo, melhorias no desempenho do sistema de treino e implementação de sistema para a melhora contínua do modelo.

Pode-se aprimorar o modelo, experimentando novas abordagens e diferentes técnicas, como diferentes configurações de autoencoders, signet, redes neurais convolucionais (CNN) e etc. Dessa forma melhorando o desempenho do sistema, uma vez que o modelo resultante seria mais assertivo.

É possível melhorar o desempenho do sistema através de paralelização e virtualização em nuvem usando técnicas de *distributed machine learning* (DML), pois, isso aumentaria o poder computacional do servidor gerando o modelo, o que abre portas para experimentos com modelos mais complexos e redes neurais mais profundas.

Outro caminho a seguir é a implementação de um modelo de *reinforcement learning*, tal técnica seria capaz de continuar aperfeiçoando o modelo enquanto esse estivesse em funcionamento.

## 3 Considerações parciais/finais

O grande foco do trabalho é comparar as técnicas de aprendizado de máquina empregadas com as metodologias clássicas de sistemas de comunicação, diferentes modulações, diferentes configurações de códigos corretores de erro, etc. E a partir desses resultados determinar se as técnicas utilizadas são comparáveis ou até melhores que as técnicas normalmente usadas.

Levando em conta a quantidade de possibilidades, o trabalho se limitará a fazer experimentos com diversas configurações de *autoencoders*, deixando para trabalhos futuros,

comparações usando redes neurais convolucionais e outras variações de redes neurais que poderiam ser utilizadas.

Do ponto de vista prático o trabalho pode ser estendido para diversos caminhos, podendo utilizar computação distribuída para melhorar o desempenho dos algoritmos, utilização de outras técnicas para codificação e decodificação da informação, explorar cenários com múltiplas antenas ou cenários ponto-multiponto. Dessa forma espera-se trazer uma variedade de trabalhos futuros.

## Referências

BENGIO, Y. et al. Learning deep architectures for ai. *Foundations and trends® in Machine Learning*, Now Publishers, Inc., v. 2, n. 1, p. 1–127, 2009.

GUPTA, A.; JHA, R. K. A survey of 5G network: Architecture and emerging technologies. *IEEE access*, IEEE, v. 3, p. 1206–1232, 2015.

HINTON, G. E. Deep belief networks. *Scholarpedia*, v. 4, n. 5, p. 5947, 2009.

JIANG, C. et al. Machine learning paradigms for next-generation wireless networks. *IEEE Wireless Communications*, IEEE, v. 24, n. 2, p. 98–105, 2017.

OSSEIRAN, A. et al. Scenarios for 5G mobile and wireless communications: The vision of the metis project. *IEEE Communications Magazine*, IEEE, v. 52, n. 5, p. 26–35, 2014.

O'SHEA, T.; HOYDIS, J. An introduction to deep learning for the physical layer. *IEEE Transactions on Cognitive Communications and Networking*, IEEE, v. 3, n. 4, p. 563–575, 2017.

RAPPAPORT, T. S. *Wireless Communications–Principles and Practice*. [S.l.]: Prentice Hall PTR, 2002.

SCHENK, T. *RF imperfections in high-rate wireless systems: Impact and digital compensation*. [S.l.]: Springer Science & Business Media, 2008.