



Universidade Federal de Campina Grande
Centro de Engenharia Elétrica e Informática
Curso de Graduação em Engenharia Elétrica

Relatório de Estágio Integrado FiberWork Optical Communications

Sílvia Claudino Martins Gomes

Campinas, SP
Agosto de 2021

Sílvia Claudino Martins Gomes

Relatório de Estágio Integrado FiberWork Optical Communications

Relatório de Estágio Integrado submetido à Coordenação de Graduação em Engenharia Elétrica da Universidade Federal de Campina Grande, campus Campina Grande, como parte dos requisitos necessários para obtenção do título de Graduado em Engenharia Elétrica.

Orientador: Edmar Candeia Gurjão

Supervisor: Euclides Lourenço Chuma

Campinas, SP
Agosto de 2021

Sílvia Claudino Martins Gomes

Relatório de Estágio Integrado FiberWork Optical Communications

Relatório de Estágio Integrado submetido à Coordenação de Graduação em Engenharia Elétrica da Universidade Federal de Campina Grande, campus Campina Grande, como parte dos requisitos necessários para obtenção do título de Graduado em Engenharia Elétrica.

Trabalho aprovado em: ____ / ____ / ____

Orientador: Prof. Dr. Edmar Candeia Gurjão

Campinas, SP
Agosto de 2021

*Dedico este trabalho aos meus pais, Bolivar e Josete, aos meus irmãos, Beatriz e Pablo,
ao meu sobrinho, Bento, e à minha avó, Lindalva (in memoriam).*

Agradecimentos

Agradeço primeiramente a minha família: aos meus pais, Josete e Bolivar, pela educação e valores transmitidos, aos meus irmãos, Beatriz e Pablo, e respectivos cunhados, Rodolfo e Natália, pelo incentivo durante o processo e ao meu sobrinho, Bento, por me ensinar mais sobre a vida do que qualquer adulto. Em especial, agradeço a minha avó, Lindalva, por ser meu exemplo de coragem e fé.

Aos meus amigos que residem em Campinas e contribuíram com o meu processo de mudança: Maria Luiza, Vidal, Machado, João, Sara, Eliseu, Rafaela e Eli. Agradeço por todo o acolhimento e por trazerem nosso Nordeste para perto.

Aos meus colegas de trabalho: Sandro, Marina, Otávio, Rafael, Talita, Danielle e Iara. A competência e compromisso de cada um torna o trabalho mais prazeroso e incentivador. Agradeço ao Dr. Sérgio por ter me concedido esta oportunidade.

Por último, agradeço ao meu supervisor, dr. Euclides, e meu orientador, professor Edmar, pela orientação durante o trabalho.

Resumo

Neste relatório estão apresentadas as atividades realizadas pela aluna de graduação em Engenharia Elétrica, Sílvia Claudino Martins Gomes, como parte da disciplina de Estágio Integrado com carga horária de 711 horas. O estágio foi realizado do período de 01 de março de 2021 à 13 de agosto de 2021 na FiberWork Optical Communications. O principal objetivo deste trabalho foi o estudo, desenvolvimento e teste de técnicas de Inteligência Artificial para identificação e classificação de dados.

Palavras-chave: Estágio Integrado, Aprendizado de Máquina, A.I., Espectroscopia.

Lista de ilustrações

Figura 1 – FiberWork Optical Communications	11
Figura 2 – Possíveis interações da radiação com a matéria	14
Figura 3 – Tela de interação com usuário	17

Sumário

1	INTRODUÇÃO	9
2	AMBIENTE DE TRABALHO	11
2.1	FiberWork Optical Communications	11
2.2	Projeto de Atuação no Estágio	12
3	FUNDAMENTAÇÃO TEÓRICA	13
3.1	Espectroscopia Vibracional	13
3.1.1	Espectroscopia Raman	14
3.1.2	Espectroscopia Infravermelho médio - MIR	14
3.2	Pré-processamento dos dados	15
3.3	Métodos de Classificação	16
4	ATIVIDADES DESENVOLVIDAS	17
4.1	Cursos Complementares	18
4.2	Demais Atividades	19
5	CONSIDERAÇÕES FINAIS	21
	REFERÊNCIAS	23

1 Introdução

A realização da disciplina de Estágio Supervisionado ou Estágio Integrado é obrigatória de acordo com o Projeto Pedagógico do Curso de Engenharia Elétrica da Universidade Federal de Campina Grande (UFCG). Com isso, o estudante torna-se apto a obter o grau de Bacharel em Engenharia Elétrica. O objetivo é que o estudante inicie no setor de atuação e, dessa forma, tenha uma experiência profissional trabalhando com aplicações dos conceitos aprendidos durante a graduação.

Este relatório tem como objetivo apresentar a experiência de estágio integrado da estudante Sílvia Claudino Martins Gomes na FiberWork Optical Communications. O período de realização foi entre 1 de março à 13 de agosto de 2021, com uma carga horária de 30 horas semanais, totalizando 711 horas, sob a orientação do professor Dr. Edmar Candeia Gurjão e a supervisão do Dr. Euclides Lourenço Chuma.

A empresa FiberWork Optical Communications desenvolve e fornece soluções de fibra óptica econômicas e tecnologicamente inovadoras. O estágio foi realizado no local da organização, no entanto, parte do trabalho necessitou ser realizado a partir do regime *Home Office* devido as medidas de proteção anunciadas pelo governo do estado de São Paulo pelo combate a pandemia do Covid-19.

Durante o estágio foram realizadas diversas atividades, dentre as quais podem ser citadas:

- Implementação de ambiente para estudo de Inteligência Artificial;
- Obtenção e criação de banco de dados para realização de estudos com Inteligência Artificial;
- Estudo das técnicas de Inteligência Artificial para identificação e classificação de dados;
- Desenvolvimento de novos Algoritmos de Inteligência Artificial para identificação e classificação;
- Realização de testes de performance dos algoritmos de Inteligência Artificial;

Este trabalho está dividido em cinco capítulos. No próximo capítulo faz-se uma descrição do grupo e da empresa onde o trabalho foi realizado. O terceiro capítulo aborda a fundamentação teórica dos conceitos utilizados. Já no Capítulo 4, encontra-se uma breve descrição das atividades desenvolvidas. No último capítulo estão as considerações finais do trabalho.

2 Ambiente de Trabalho

2.1 FiberWork Optical Communications

Fundada em 1998 por Dr. Sérgio Barcellos, a FiberWork desenvolve e fornece soluções de fibra óptica econômicas e tecnologicamente inovadoras para os mercados de telecomunicações e não-telecomunicações. Entre os serviços oferecidos pela empresa, destaca-se: contrato de pesquisa e desenvolvimento; consultoria; treinamentos e cursos técnicos; planejamento, especificação, projeto, instalação, testes e diagnósticos de redes de fibra óptica; além de suporte técnico em suas áreas de atuação. [1]

Figura 1 – FiberWork Optical Communications



Fonte: Autoria Própria

A empresa Fiberwork está localizada no primeiro andar do edifício apresentado na Figura 1. Já o Instituto de Inovações Fotônicas (ITech) situa-se no térreo do mesmo prédio. Ao descrever as atividades realizadas, será possível expor a colaboração entre as entidades.

2.2 Projeto de Atuação no Estágio

Devido às políticas de privacidade da Fiberwork, o projeto no qual foram realizadas as atividades não poderá ser nomeado diretamente. Além disso, a base de dados obtida com o estudo descrito é confidencial.

Na empresa, existem diversos projetos que são desenvolvidos concomitantemente por equipes distintas. A estagiária integra uma das equipes composta por cinco pessoas: uma química, uma bióloga, um estagiário em química, uma estagiária em engenharia elétrica e um desenvolvedor. Ocorrem reuniões semanais com a equipe citada e o gerente de Pesquisa e Desenvolvimento para relatar as atividades executadas no período.

3 Fundamentação Teórica

Neste capítulo serão abordados os conceitos de espectroscopia, pré-processamento de dados e aprendizado de máquina que foram utilizados para obtenção, preparação e análise da base de dados.

3.1 Espectroscopia Vibracional

Espectroscopia vibracional é o estudo da interação da matéria com a radiação eletromagnética. A interação pode causar absorção, emissão ou reflexão, a depender da natureza da fonte e da própria amostra. Com isso, é possível identificar substâncias químicas ou grupos funcionais nas formas sólida, líquida ou gasosa. [2]

A espectroscopia vibracional permite a análise rápida com pouco ou sem preparo de amostra e medidas de forma não destrutiva. A espectroscopia na região do infravermelho pode ser subdividido em: infravermelho próximo (NIR: *Near Infrared*), infravermelho médio (MIR, *Mid Infrared*) e infravermelho distante (FIR: *Far infrared*). A tabela a seguir apresenta as faixas espectrais em comprimento de onda e número de onda para cada técnica citada.

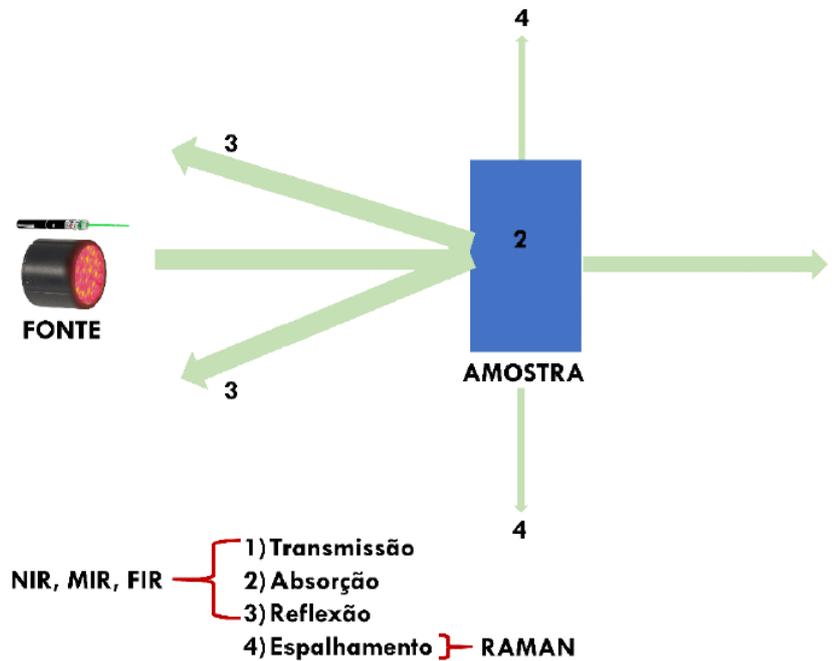
Tabela 1 – Faixas espectrais para as diferentes regiões do infravermelho: Próximo, Médio e Distante

Região	Comprimento de onda (nm)	Número de onda (cm ⁻¹)
Próximo	780-2500	12800-4000
Médio	2500-5000	4000-200
Distante	5000-10000	200-10

Fonte: Neves, Marina de Gea (2021, p. 35)

Além das técnicas NIR, MIR e FIR, existe a espectroscopia Raman que baseia-se no espalhamento da radiação. A figura 1 apresenta os efeitos da interação entre radiação e amostra observados na espectroscopia vibracional. Neste trabalho foram analisadas bases de dados provenientes das técnicas Raman e MIR, com isso, os demais métodos não serão abordados em detalhes.

Figura 2 – Possíveis interações da radiação com a matéria



Fonte: Neves, Marina de Gea (2021, p. 36) [2]

3.1.1 Espectroscopia Raman

O fenômeno de espalhamento pode ocorrer, basicamente, por meio de dois processos: elástico e inelástico. O primeiro, que ocorre de maneira majoritária, também conhecido como espalhamento Rayleigh, se caracteriza pela energia dos fótons incidentes ser a mesma dos fótons espalhados, não fornecendo informações úteis ao sistema. O segundo processo, conhecido por espalhamento Raman, ocorre naturalmente 0,0000001%, sendo caracterizado pela diferença de energia dos fótons incidentes e dos espalhados.

Esse fenômeno é observado somente quando as transições vibracionais envolvidas ocorrem com a variação da polarizabilidade da molécula. Portanto, depende da estrutura química do analito. Logo, cada banda representado no espectro Raman corresponde a uma vibração de ligação molecular específica.

3.1.2 Espectroscopia Infravermelho médio - MIR

A espectroscopia do infravermelho médio identifica estruturas químicas com base na absorção da luz infravermelha incidida sobre a amostra. Para que ocorra o fenômeno da absorção, é necessário que a molécula sofra variação no momento dipolo. Se o momento de dipolo for nulo, seu espectro vibracional não será observado no infravermelho, somente no

Raman, caso possua modos ativos na técnica. A explicação se dá pelo fato da variação do momento dipolo, produzir um campo que pode interagir com o campo elétrico da radiação incidente

Uma das técnicas mais utilizadas atualmente na espectroscopia MIR é a ATR-FTIR (Reflectância Total Atenuada-Espectroscopia de infravermelho com transformada de Fourier). Nessa técnica, uma radiação na região do infravermelho médio é incidida através de um cristal sobre uma amostra depositada em sua superfície. Em seguida, parte da radiação é absorvida pela amostra. A radiação remanescente é refletida para o detector tornando possível analisar as interações.

3.2 Pré-processamento dos dados

Uma etapa fundamental para classificação dos sinais espectrais é o tratamento. Ao analisar o espectro bruto, existem fatores no processo de aquisição dos dados que interferem a leitura do sinal. Ruído, variação de temperatura da amostra e o estado da matéria podem causar prejuízo no reconhecimento da representação espectral. Com isso, utiliza-se diversas técnicas de pré-processamento para recuperação.

O filtro Savitzky-Golay tem como objetivo suavizar os dados, tornando possível aumentar a precisão dos dados espectrais sem distorcer a tendência do sinal. O filtro utiliza o método dos mínimos quadrados lineares para ajustar subconjuntos sucessivos de pontos de dados adjacentes com um polinômio de baixo grau.

Outra técnica que está sendo amplamente utilizada nas aplicações espectrais é denominada Mínimos Quadrados Penalizados ponderados iterativamente adaptativos (airPLS). O método funciona alterando iterativamente os pesos dos erros de soma dos quadrados (SSE) entre a linha de base ajustada e os sinais originais. Com isso, os pesos do SSE são obtidos de forma adaptativa usando a diferença entre a linha de base ajustada anteriormente e os sinais originais. [3]

Normalização e corte são etapas simples mas fundamentais no tratamento dos dados, minimizando sinais não relevantes as informações químicas. A normalização é um método que garante a integridade dos dados. Já o corte do intervalo desejado torna possível remover partes espectrais não relevantes e/ou regiões com alta concentração de ruídos.

3.3 Métodos de Classificação

Os métodos de Algoritmo Genético aliado a Análise de Discriminante Linear (GA-LDA) e Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA) são métodos de classificação utilizados em problemas de quimiometria.

Algoritmos genéticos são uma classe de técnicas de otimização. Inspirado na teoria evolutiva que considera fatores evolucionários como hereditariedade, mutação, seleção natural e recombinação para realizar o procedimento. O GA é capaz, portanto, de auxiliar na seleção de variáveis.

O método LDA é usado como uma ferramenta para classificação, redução de dimensão e visualização de dados. O método é uma generalização do discriminante linear de Fisher, um método usado em estatísticas e outros campos, para encontrar uma combinação linear de recursos que caracterizam ou separam duas ou mais classes de objetos ou eventos. [4] No entanto, o método possui dificuldade ao se trabalhar com dados que disponham de um quantidade elevada de variáveis e pouca informação útil para distinção de classes. Sendo a espectroscopia uma técnica de análise multivariada, seus dados resultam em grande quantidade de variáveis, sendo necessário a associação com métodos de redução, como por exemplo, o algoritmo genético.

O PLS-DA combina redução da dimensionalidade e análise discriminante em um algoritmo. O método executa uma regressão multivariada projetando os dados em um espaço de dimensão inferior, chamado espaço das variáveis latentes. Neste novo espaço é possível calcular a variância acumulada entre a matriz de entrada (X) e a matriz categórica (Y) utilizando diferentes números de variáveis latentes. Esse número deve ser otimizado a fim de maximizar a covariância e evitar o sobreajuste (*overfitting*) do modelo. Ao final, a previsão é realizada com valores contínuos, visto que é construído um limiar (*threshold*) que realiza a distinção das classes a partir dos valores obtidos pela regressão.

4 Atividades Desenvolvidas

A estagiária realizou as atividades de aplicações de Inteligência Artificial no projeto. Para isso, o primeiro mês de trabalho foi dedicado ao estudo das técnicas de aprendizado de máquina para identificação e classificação de dados. Com isso, foi possível selecionar os principais métodos utilizados para o problema proposto.

Após essa etapa, a estagiária buscou na literatura base de dados para a realização de testes com os métodos selecionados. Desse modo, foi possível criar um banco de dados para realização de estudos com Inteligência Artificial. Com isso, o trabalho foi otimizado, visto que antes da aquisição dos dados em laboratório, já havia um estudo dos métodos mais eficazes para tratamento e classificação dos sinais espectrais.

Além disso, foi desenvolvido um algoritmo na linguagem Python de interação com o usuário para a implementação da base de dados estudada. O usuário seleciona os métodos de pré-processamento e os respectivos parâmetros, como é possível visualizar na figura 3. Ao final, o programa retorna os valores de acurácia, precisão, sensibilidade e especificidade da validação cruzada e teste. Além de gráficos contendo a função de decisão e gráfico de resíduos.

Figura 3 – Tela de interação com usuário

```

model = GaLDA.model_GaLDA(df)
Name of class 1: Classe 1
Name of class 2: Classe 2
Customize preprocessing? (y/n) y
Choice the methods
Available methods: savgol, airPLS and normalize
Select 1º method (enter end to finalize choice): savgol
Select 2º method (enter end to finalize choice): airPLS
Select 3º method (enter end to finalize choice): end
The methods to be applied will be: ['savgol', 'airpls']
Confirm? (y/n) y
Customize Savtzyk-golay? (y/n) y
window lenght (standard = 9): 7
polynomial order (standard = 2): 4
derivative order (standard = 0): 0
savitzky-golay filter applied.
Customize airPLS? (y/n) n
4% | ██████████
7, 1.68it/s] | 11/309 [00:06<02:5

```

Fonte: Autoria Própria

Os modelos desenvolvidos para a base de dados obtida em laboratório obteve resultados satisfatórios. A tabela a seguir apresenta os valores dos parâmetros para os métodos GA-LDA e PLSDA.

Tabela 2 – Resultado do teste para os métodos PLSDA e GA-LDA.

	PLSDA	GA-LDA
Acurácia	92%	96%
Precisão	87%	95%
Sensibilidade	95%	95%
Especificidade	90%	97%
F1-Score	91%	95%
NPV	96%	97%

Fonte: Autoria Própria

4.1 Cursos Complementares

O projeto possui o apoio financeiro da Financiadora de Estudos e Projetos (FINEP), empresa pública do Ministério da Ciência, Tecnologia e Inovações (MCTI). O órgão incentiva que os participantes dos projetos realizem cursos complementares que contribuam para o desenvolvimento do projeto. A estagiária recebeu permissão para eleger cursos e apresentar à coordenadora justificando como o assunto tratado seria útil para o trabalho na empresa.

A estagiária concluiu duas formações. A primeira é intitulada "Neural Networks and Deep Learning" autorizado pelo DeepLearning.AI, com carga horária de 23 horas. Em seguida, a estudante participou do curso "Databases and SQL for Data Science with Python" autorizado pela International Business Machines Corporation (IBM). A formação possui uma carga horária de 18 horas. Ambos os cursos são oferecidos pela plataforma Coursera.

Além disso, foi realizado um treinamento em técnicas de vendas e negociação. A funcionária de marketing da FiberWork ministrou o curso. O objetivo é desenvolver habilidades de comunicação e influência para os participantes terem mais sucesso ao expor propostas e ideias.

Por último, foi criado um grupo de estudos em Inteligência Artificial composto pelos quatro colaboradores que trabalham exclusivamente com I.A. na FiberWork e na Itech. O objetivo do grupo é o compartilhamento de conhecimento a fim de promover uma maior eficiência nos projetos.

4.2 Demais Atividades

Outra atividade pertinente ao projeto foi o desenvolvimento do aplicativo mobile na App Store. Ocorreram reuniões semanais, juntamente com os desenvolvedores de software, com o intuito de discutir e implementar ideias. Ficou a encargo da estagiária, reunir feedbacks de sete usuários de testagem e apresentá-los nas reuniões. A cada nova atualização e rodada de testes, as considerações eram discutidas e votadas para novas modificações no aplicativo.

5 Considerações Finais

Neste trabalho foram descritas as atividades realizadas pela autora durante o estágio. Nesse período, foi possível colaborar com o conhecimento em Inteligência Artificial para o projeto na FiberWork Optical Communications. A experiência foi extremamente enriquecedora, pois a estagiária estava inserida em uma equipe multidisciplinar e o ambiente era de colaboração e aprendizado.

A Inteligência Artificial é um campo novo da ciência e, por isso, ainda não há disciplinas no curso de Engenharia Elétrica que abordam o assunto. No entanto, a base dos métodos de IA e os tratamentos de sinais espectrais foram estudados pela estagiária nas disciplinas de Álgebra Linear, Análise de Sinais e Sistemas e Processamento Digital de Sinais. Dessa forma, a aluna pôde colocar em prática muitos conceitos adquiridos durante a graduação.

Por fim, a aluna pôde desenvolver habilidades técnicas e subjetivas. A confiança e autonomia dada pela empresa à estagiária colaborou com essa evolução. Com isso, foi possível desenvolver as competências de: inteligência emocional, resiliência, comunicação, flexibilidade e trabalho em equipe. Além disso, foi possível compreender e vivenciar as diversas etapas de um projeto.

Referências

- [1] FiberWork Optical Communications. Disponível em <www.fiberwork.net/> Acesso em agosto de 2021. Citado na página 11.
- [2] M. de Gea Neves. Avaliação de diferentes métodos quimiométricos para autenticação, classificação e quantificação em matrizes alimentícias empregando espectroscopiavibracional. Campinas – SP. 2021. Citado 2 vezes nas páginas 13 e 14.
- [3] Zhang, Zhimin; Liang, Yi-Zeng. **Baseline correction using adaptive iteratively reweighted penalized least squares**. PubMed, 2010. Citado na página 15.
- [4] *Linear discriminant analysis* Disponível em <https://en.wikipedia.org/wiki/LineardiscriminantanalysisLDAfortwoclasses/> Acesso em agosto de 2021. Citado na página 16.
- [5] Espectroscopia. Disponível em <<https://www.infoescola.com/fisica/espectroscopia/>> Acesso em agosto de 2021. Nenhuma citação no texto.
- [6] Reflectância Total Atenuada (ATR). Disponível em <<https://www.mt.com/br/pt/home/products/L1AutochemProducts/ReactIR/attenuated-total-reflectance-atr.html>> Acesso em agosto de 2021. Nenhuma citação no texto.
- [7] Espectroscopia Raman. Disponível em <<https://www.mt.com/br/pt/home/applications/L1AutoChemApplications/Raman-Spectroscopy.html>> Acesso em agosto de 2021. Nenhuma citação no texto.
- [8] *What is Raman Spectroscopy?* Disponível em <<https://www.horiba.com/en/en/raman-imaging-and-spectroscopy/>> Acesso em agosto de 2021. Nenhuma citação no texto.
- [9] Lee, Loong Chuen; Liang, Choong-Yean; Jemain, Abdul Aziz. **Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps**. Analyst, 2018. Nenhuma citação no texto.

- [10] *Linear Discriminant Analysis, Explained* Disponível em <https://towardsdatascience.com/linear-discriminant-analysis-explained-f88be6c1e00b> Acesso em agosto de 2021. Nenhuma citação no texto.
- [11] *Linear Discriminant Analysis for Machine Learning* Disponível em <https://machinelearningmastery.com/linear-discriminant-analysis-for-machine-learning/> Acesso em agosto de 2021. Nenhuma citação no texto.
- [12] Cui, Minshan; Prasad, Saurabh; Mahroogh, Majid; Bruce, Lori M.; Aanstoos, James. **Genetic Algorithms and Linear Discriminant Analysis Based Dimensionality Reduction Remotely Sensed Image Analysis**. IGARSS, 2011. Nenhuma citação no texto.