



PROJETO PEDAGÓGICO DE CURSO
Pós-Graduação *Lato Sensu*
Engenharia de dados

Sumário

1. Nome do Curso e Área do Conhecimento.....	3
2. Características Técnicas do Curso.....	3
3. Público-Alvo	3
4. Critérios de Seleção	4
5. Justificativa do Curso	4
6. Objetivos do Curso	4
7. Competências e Habilidades do Curso	5
8. Metodologia de Ensino e Aprendizagem	5
9. Estágio Não Obrigatório.....	8
10. Matriz Curricular	10
11. Carga Horária.....	11
12. Conteúdo Programático.....	11
13. Infraestrutura Física e Pedagógica.....	20

1. Nome do Curso e Área do Conhecimento

Nome do Curso: Engenharia de dados

Área de Avaliação (CAPES): CIÊNCIA DA COMPUTAÇÃO

Grande Área (CAPES): 10000003 CIÊNCIAS EXATAS E DA TERRA

Área do Conhecimento (CAPES): 10300007 CIÊNCIA DA COMPUTAÇÃO

Classificação OCDE: Computação e Tecnologias da Informação e Comunicação (TIC)

2. Características Técnicas do Curso

Modalidade: Educação a Distância

Número máximo de vagas por Polo/Unidade: 1000 alunos

Período de Oferecimento: O curso possui entrada intermitente, respeitadas as datas de início e de fim cadastradas na oferta, bem como observado o período indicado para a sua integralização.

Limitações legais

Resolução CNE/CES Nº 1, de 06 de abril de 2018, que estabelece normas para o funcionamento de cursos de pós-graduação lato sensu.

O candidato deverá ser graduado com diploma devidamente registrado segundo as normas estabelecidas pelo MEC.

3. Público-alvo

Graduados em cursos de tecnologia ou engenharia, profissionais de mercado atuantes em tecnologia da informação, administradores de redes, analistas de dados, analistas de business intelligence e cientistas de dados.

4. Critérios de Seleção

O ingresso na pós-graduação será realizado por meio de inscrição no portal da Instituição e entrega dos documentos solicitados. Em seguida, analisados pela área competente.

5. Justificativa do Curso

O aumento no volume de dados, as diferentes formas de trabalhar a informação e a crescente utilização de tecnologias para suportar essas novas modalidades de lidar com a informação tornaram elevada a procura por um perfil profissional capaz de coletar, processar, armazenar e distribuir dados. Utilizando ferramentas para projetar, implementar, construir e integrar fontes de informação de diferentes origens, o Engenheiro de Dados elabora soluções técnicas orientado pela visão do negócio em distribuir a informação, criando soluções de ponta a ponta que reflitam o pensamento estratégico da organização.

6. Objetivos do Curso

6.1. Objetivo Geral

- Possibilitar que o aluno atue como Engenheiro de Dados, coletando, transformando, armazenando, gerenciando e disponibilizando dados corporativos às organizações e agentes envolvidos, participando também do processo de tomada de decisão.

6.2. Objetivos Específicos

- Oferecer formação específica na área de ciência e engenharia de dados de maneira multidisciplinar, difundindo e disseminando o conhecimento dos fundamentos e da atividade prática;

- Contribuir com uma formação prática por meio de ferramentas atuais e presentes no mercado;
- Capacitar na identificação e proposição de soluções técnicas aos problemas das organizações e da sociedade; e
- Fornecer uma visão estratégia dos dados nas organizações.

7. Competências e Habilidades do Egresso

Competências

Atuar como Engenheiro de Dados, coletando, transformando, armazenando, gerenciando e disponibilizando os dados a agentes corporativos, internos e externos, de forma a manter a integridade da informação e auxiliar no processo de tomada de decisão.

Habilidades

- Atuar em conjunto com DBAs (Administradores de Banco de Dados) e Analistas de Dados, visando à integração de sistemas computacionais, de ponta a ponta;
- Programar em linguagens de banco de dados relacionais, como PL/SQL e T-SQL, e específicas para Big Data, como Python;
- Auxiliar tomadores de decisão através da construção artefatos OLAP provenientes de Data Warehouse e Data Marts;
- Construir Data Warehouse utilizando programação e ferramentas específicas para integração de dados;
- Criar, administrar e compreender aspectos relacionados a infraestrutura tecnológica responsável pela sustentação dos bancos de dados e aplicações relacionadas.

8. Metodologia de Ensino e Aprendizagem

O desenvolvimento das disciplinas do curso se dará no ambiente virtual, onde o aluno cumprirá 40 horas por disciplina.

Estas 40 horas são compostas pelo estudo dos diversos insumos pedagógicos disponibilizados, como:

- ✓ materiais de leitura;
- ✓ videoaulas;
- ✓ slides;
- ✓ podcasts;
- ✓ indicações de leituras extras como artigos e capítulos de livros.

Como parte do modelo acadêmico, também contabilizamos:

- ✓ a interação com os tutores para esclarecimentos de dúvidas pedagógicas (via Fórum);
- ✓ a realização da avaliação disciplinar pelo aluno.

Vale ressaltar que cada aluno é único e tem seu próprio tempo para aprendizagem. Alguns conseguem assimilar o conteúdo com mais agilidade, outros precisam retomar as leituras e videoaulas para tornar sua aprendizagem efetiva.

O tempo gasto pelo aluno para o entendimento de cada disciplina depende também da sua familiaridade com o tema, ou seja, o conhecimento prévio que o aluno tem em relação a temática abordada.

No ambiente virtual, o aluno encontrará o conteúdo das disciplinas, organizados em temas/webaulas.

Para cada um deles, o aluno realizará um conjunto de atividades baseadas em leitura de textos de fundamentação teórica e acesso a recursos audiovisuais.

Um tutor apoiará as atividades realizadas no ambiente virtual, atendendo o aluno nas suas dúvidas por meio de ferramentas de comunicação.

O aluno, ao iniciar os seus estudos, terá um encontro presencial para acolhida/ambientação; esse encontro terá como objetivos:

- ✓ Integrar o aluno ao curso de Pós-Graduação.
- ✓ Dialogar e esclarecer as dúvidas sobre a proposta pedagógica do curso e as regras acadêmicas.
- ✓ Apresentar ao aluno o Ambiente Virtual de Aprendizagem (o primeiro acesso; o envio de documentos; os serviços de secretaria e financeiro; a disciplina Ambientação; a tutoria online; o boletim acadêmico; as disciplinas e seus conteúdos; a biblioteca virtual; entre outros).
- ✓ Proporcionar um momento de Network aos pós-graduandos.

Avaliação do Desempenho do Aluno

O aluno deverá realizar as atividades propostas no ambiente virtual. A realização das atividades irá compor sua frequência no curso, que será considerada para a sua aprovação.

A atividade avaliativa que o aluno realizará para compor a sua média é a Avaliação Virtual (AV); essa atividade é obrigatória e estará disponível no Ambiente Virtual de Aprendizagem – AVA, conforme cronograma de seu curso.

Para a aprovação em cada uma das disciplinas, o aluno deverá obter frequência de, no mínimo, 75% (setenta e cinco por cento) e nota igual ou superior a 7,0 (sete).

As notas devem ser expressas no intervalo de 0 (zero) a 10 (dez).

O aluno que obtiver média inferior a 7,0 (sete) nas disciplinas terá direito ao Programa de Dependência e Recuperação – PDR, mediante a solicitação de requerimento e respeitando o período de jubramento do curso.

O PDR será realizado no ambiente virtual de aprendizagem, sendo que o aluno terá acesso ao conteúdo da disciplina e realizará uma Avaliação Virtual - AV, e a nota obtida substituirá a média do aluno.

Para a obtenção do **Certificado** de Pós-graduação *Lato Sensu* – especialização, o aluno deverá cumprir todas as condições seguintes:

- ✓ Frequência mínima de 75% (setenta e cinco por cento) em todas as disciplinas;
- ✓ Nota igual ou superior a 7,0 (sete) em todas as disciplinas.

Certificação

O Certificado de conclusão de curso de Especialização será acompanhado por histórico escolar, em cumprimento às exigências da Resolução CNE/CES nº1, de 06 de abril de 2018, da Câmara de Educação Superior do Conselho Nacional de Educação.

Composição do Corpo Docente

O corpo docente do curso é constituído por profissionais qualificados, com comprovado saber em sua área de atuação, conforme Resolução CNE/CES nº1, de 06 de abril de 2018, sendo integrado, no mínimo, por 30% (trinta por cento) de portadores de título de pós-graduação *stricto sensu*, isto é, portadores de títulos de Mestrado e Doutorado, obtidos em programas de pós-graduação *stricto sensu* devidamente reconhecidos pelo poder público em território nacional, ou revalidados, conforme legislação vigente. Os demais docentes são certificados em nível de especialização, pós-graduação *lato sensu*, de reconhecida capacidade técnico-profissional.

9. Estágio Não Obrigatório

O estágio curricular não obrigatório tem como finalidade estimular o aluno a desenvolver atividades extracurriculares, para que possa inter-relacionar os conhecimentos teóricos e práticos adquiridos durante o curso e aplicá-los na solução de problemas reais da profissão, proporcionando o desenvolvimento da análise crítica e reflexiva para os problemas socioeconômicos do país, de acordo com a Resolução de Estágio curricular não obrigatório vigente na Instituição.

O estágio curricular não obrigatório do curso de Pós-Graduação *Lato Sensu* segue as diretrizes estabelecidas pela Lei Nº 11.788/2008.

Os principais objetivos da prática do estágio curricular não obrigatório são:

I. proporcionar o exercício do aprendizado comprometido com a realidade socioeconômica-política do país;

II. propiciar a realização de experiências de ensino e aprendizagem visando à educação profissional continuada, alicerçada no desenvolvimento de competências e habilidades e ao exercício do pensamento reflexivo e criativo; e

II. incentivar o trabalho de pesquisa e investigação científica, visando ao desenvolvimento da ciência, da tecnologia e da cultura.

A carga horária é definida pela concedente de estágio, não podendo ultrapassar a carga horária máxima de 6 (seis) horas diárias e 30 (trinta) horas semanais, as quais podem ser realizadas em empresas públicas ou privadas, instituição de pesquisa, órgãos governamentais e não governamentais, e as próprias unidades da Universidade, desde que sigam às condições adequadas para que o estagiário possa aprofundar os seus conhecimentos teóricos e práticos adquiridos no curso.

Para o Curso de Pós-Graduação EAD, a prática do estágio curricular não obrigatório é permitida durante a vigência do curso, não podendo exceder em um mesmo campo de estágio o período de 2 (dois) anos.

Os estágios curriculares não obrigatórios devem estar apoiados em Termo de Compromisso e de comum acordo com a Instituição, devendo explicitar não somente os aspectos legais específicos, como também os aspectos educacionais e de compromisso com a realidade social.

O Planejamento do Estágio Curricular Não Obrigatório é de responsabilidade do coordenador de curso/professor orientador e também do Departamento de Estágios.

10. Matriz Curricular

DISCIPLINAS	CH PRÁTICA	CH TEÓRICA	CH TOTAL
Ambientação	0 h	0 h	0 h
Modelagem e arquitetura do dw (data warehouse)	0 h	40 h	40 h
Interações entre big data e cloud computing	0 h	40 h	40 h
Banco de dados relacional e big data	0 h	40 h	40 h
Governança de dados	0 h	40 h	40 h

Bancos de dados não relacionais (NoSQL)	0 h	40 h	40 h
Machine learning	0 h	40 h	40 h
Projeto em ciência de dados com soluções para processamento paralelo e distribuído de dados	0 h	40 h	40 h
Administração de bancos de dados	0 h	40 h	40 h
Gerenciamento do desempenho do banco de dados (tuning)	0 h	40 h	40 h

11. Carga Horária

A carga horária de 360h constitui o conteúdo ministrado em 9 (nove) disciplinas.

12. Conteúdo programático

Disciplina: Ambientação

Ementa: Histórico da Educação a Distância. Legislação da Educação a Distância no Brasil. Potencialidades da Educação a Distância. Flexibilidade de Acesso. Tecnologias para apoio à pesquisa. Aprendizagem colaborativa. Características do aluno na EAD. Boa convivência virtual: netiquetas.

Competências e Habilidades:

- Identificar as características do aluno da modalidade de educação a distância;
- Utilizar ferramentas tecnológicas de apoio a pesquisa;
- Comunicar-se com clareza e assertividade oralmente e por escrito;
- Gerenciar tempo e atividades de estudo.

Conteúdo Programático 1: Histórico da Educação a Distância.

Conteúdo Programático 2: Legislação da Educação a Distância no Brasil.

Conteúdo Programático 3: Potencialidades da Educação a Distância.

Conteúdo Programático 4: Flexibilidade de Acesso.

Conteúdo Programático 5: Tecnologias para apoio à pesquisa.

Conteúdo Programático 6: Aprendizagem colaborativa.

Conteúdo Programático 7: Características do aluno na EAD.

Conteúdo Programático 8: Boa convivência virtual: netiquetas.

Bibliografia básica:

GOTTARDI, M. de L. A autonomia na aprendizagem em educação a distância: competência a ser desenvolvida pelo aluno. **Associação Brasileira de Educação a Distância**, São Paulo, v. 14, n. 8, p. 110-123, dez, 2015.

MORAN, J. M. MASETTO, M. T. BEHRENS, M. A. **Novas Tecnologias e Mediação Pedagógica**. 21. ed. Campinas, SP: Papirus, 2013.

PIVA, D. J. PUPO, R. GAMEZ, L. OLIVEIRA, S. **EAD na Prática: Planejamento, métodos e ambientes de educação online**. São Paulo: Elsevier, 2011.

Bibliografia Complementar:

ALVES, L. Educação a distância: conceitos e história no Brasil e no mundo. **Associação Brasileira de Educação A Distância**, São Paulo, v. 10, n. 7, p.85-92, out. 2011.

MAIA, Carmen; MATTAR, João. **ABC da EaD: a educação a distância hoje**. São Paulo: Prentice Hall, 2007.

MOORE, Michael G., **Educação a distância: sistemas de aprendizagem on-line**. 3. ed., São Paulo: Cengage Learning, 2013.

Disciplina: Modelagem e arquitetura do DW (data warehouse).

Ementa: A disciplina de Modelagem e Arquitetura do DW (Data Warehouse) traz conhecimentos de modelos e arquiteturas condizentes com DW. Mostra ferramentas construção e verificar dados dentro das estruturas.

Competências e Habilidades:

- Reconhecer e modelar a arquitetura DW;
- Reconhecer e utilizar ferramentas em DW;
- Efetuar a obtenção de dados dentro do ambiente.

Conteúdo Programático 1: Banco de Dados Transacionais X Bancos de Dados Analíticos.

Conteúdo Programático 2: Conceitos básicos sobre Data Warehouse.

Conteúdo Programático 3: Data Mart.

Conteúdo Programático 4: Modelagem de Dados para um Data Warehouse.

Conteúdo Programático 5: Esquema Estrela e Esquema Floco de Neve.

Conteúdo Programático 6: Ferramentas de Dados em um Data Warehouse.

Conteúdo Programático 7: Mineração de Dados em Data Warehouse.

Bibliografia Básica:

INMON, W. H. **Como Construir o Data Warehouse**. Rio de Janeiro: Campus, 1997.

KIMBALL, R. **The Data Warehouse Toolkit**: guia completo para modelagem dimensional. Rio de Janeiro: Campus, 2002.

Machado, F. N. **Tecnologia e Projeto de Data Warehouse**: uma visão multidimensional. São Paulo: Érica, 2004.

KIMBALL, Ralph. **Data Warehouse – Construindo o Data Warehouse para Web**. Rio de Janeiro: Campus, 2000.

Bibliografia Complementar:

RASLAN, D. A.; CALAZANS, Angélica Toffano Seidel. **Data Warehouse: Conceitos e Aplicações**. Brasília, DF, Abril de 2014. Disponível em: <<https://www.publicacoesacademicas.uniceub.br/gti/article/viewFile/2612/2400>> . Acesso em 25 set. 2018.

ZIULKOSKI, L. C. C. **Coleta de Requisitos e Modelagem de Dados para Data Warehouse**: Um Estudo de Caso utilizando Técnicas de Aquisição de Conhecimento. Porto Alegre, RS, Março de 2003. Disponível em: <<http://www.inf.ufrgs.br/bdi/wp-content/uploads/Monografia-LuisZiulkoski.pdf>>. Acesso em 25 set. 2018.

NOVAIS, R. C. N. **Modelagem Dimensional**. São Paulo, SP, Agosto de 2012 .Disponível em: <<http://www.fatecsp.br/dti/tcc/tcc00071.pdf>>. Acesso em 25 set. 2018.

Disciplina: Interações entre big data e cloud computing.

Ementa: Introdução ao Big Data: Fundamentos, Infraestrutura e Interfaces. Gerenciamento de Big Data: Coleta e Processamento de Dados. Aplicação da Ciência de Dados no gerenciamento empresarial. Algoritmos de Aprendizado de Máquina para minerar os dados. Big Data Analytics. Estruturas de Programação em Nuvem. Cloud Computing e Big Data.

Competências e Habilidades:

- Desenvolver a habilidade de raciocinar em ambientes computacionais de alta complexidade.
- Compreender os fundamentos da análise, gerenciamento e armazenamento de dados.

- Assimilar a importância do Aprendizado de Máquina na mineração de dados.
- Analisar a relação entre Cloud Computing e Big Data.

Conteúdo Programático 1: Introdução ao Big Data: Fundamentos, Infraestrutura e Interfaces.

Conteúdo Programático 2: Gerenciamento de Big Data: Coleta e Processamento de Dados.

Conteúdo Programático 3: Aplicação da Ciência de Dados no gerenciamento empresarial.

Conteúdo Programático 4: Big Data Analytics.

Conteúdo Programático 5: Algoritmos de Aprendizado de Máquina para minerar os dados.

Conteúdo Programático 6: Cloud Computing e Big Data.

Conteúdo Programático 7: Estruturas de Programação em Nuvem.

Bibliografia Básica:

MACHADO, F. N. R. **Big data:** o futuro dos dados e aplicações. São Paulo: Editora Erica, 2018

KELLEHER, J. D.; TIERNEY, B. **Data science.** Cambridge: MIT Press Essential Knowledge Series, 2018

RUPARELIA, N. B. **Cloud computing.** Cambridge: MIT Press Essential Knowledge Series, 2016

NETO, M. V. S. **Computação em nuvem.** Rio de Janeiro: Editora Brasport, 2015

SAGIROGLU, S.; SINANC, D. **Big data:** a review. International Conference on Collaboration Technologies and Systems (2013), 13681548.

HASHEM, I. A. T.; YAQOOB, I.; ANUAR, N.B.; MOKHTAR, S.; GANI, A.; KHAN, S.U. **The rise of “big data” on cloud computing**: review and open research issues. Information Systems 47 (2015), 98 - 115.

QI, H.; GANI, A. **Research on mobile cloud computing**: Review, trend and perspectives. Second International Conference on Digital Information and Communication Technology and it's Applications (2012), 12804207.

JULA, A.; SUNDARARAJAN, E.; OTHMAN, Z. **Cloud computing service composition**: A systematic literature review. Expert Systems with Applications 41 (2014), p. 3809-3824.

Bibliografia Complementar:

Data Science and Big Data Analytics: Making Data-Driven Decisions. Disponível em: <<https://bigdataanalytics.mit.edu>> Acesso em 13 nov.2018.

MIT-IBM Watson AI Lab. Disponível em: <<https://mitibmwatsonailab.mit.edu>> Acesso em 13 nov.2018. What is Cloud Computing? Disponível em: <<https://www.ibm.com/cloud/learn/what-is-cloud-computing>> Acesso em 13 nov.2018.

Nebula Cloud Computing Platform. Disponível em: <<https://www.nasa.gov/open/nebula.html>> Acesso em 13 nov.2018.

Disciplina: Banco de dados relacional e big data.

Ementa: Estruturação do Banco de Dados Relacional, Modelos de Entidade e Relacionamento, Conceitos de transações, Blocos de Código; Princípios da Linguagem SQL e seus Comandos. Introdução aos Sistemas de Gestão de Banco de Dados, bem como ao Big Data, repositórios de dados, Data Lake, e armazenamento em nuvem.

Competências e Habilidades:

- Compreender a importância do armazenamento de dados de forma adequada no modelo relacional do banco de dados;
- Estudar possíveis aplicações do modelo relacional do banco de dados;
- Aprender os princípios básicos da linguagem SQL;
- Entender as razões que calçam a crescente utilização do Big Data.

Conteúdo Programático 1: Banco de Dados e Big Data: União Necessária. Modelo Entidade Relacionamento; Estrutura de Banco de Dados Relacional; Modelo Relacional; Bloqueio e Simultaneidade de Banco de Dados; Banco de Dados Relacional do Futuro: relacional e autônomo; introdução ao Big Data; ao Repositório de Dados; ao Data Lake e ao armazenamento em Nuvem.

Conteúdo Programático 2: ACID e Benefícios do Banco de Dados Relacional. Princípios de Atomicidade; Consistência de Dados; Integridade e Durabilidade do banco de dados Relacional; Conceito de transações; Cardinalidade entre as transações (cardinalidades máximas e mínimas); Benefícios dos Bancos de Dados Relacionais; Aplicações e Futuro do Banco de Dados Relacional.

Conteúdo Programático 3: A SQL e os Sistemas de Gestão de Banco de Dados. Compreensão das características básicas da Linguagem Estruturada de Pesquisa (SQL); principais comandos embasados em DDL e DML; vantagens e desvantagens do SQL; introdução aos Sistemas de Gestão de Banco de Dados no meio empresarial e a relação com a SQL.

Conteúdo Programático 4: Big Data e bancos de dados não relacionais. Compreendendo o conceito e aplicabilidade da Big Data; 5 “V” do Big Data; ferramentas; Cenário atual; gerenciamento de dados com Hadoop e NoSQL.

Bibliografia:

- COSTA, R. L. **SQL: Guia Prático**. 2. ed. Rio de Janeiro: Brasport, 2006.
- DATE, C. J. **SQL e Teoria Relacional: Como escrever códigos SQL precisos**. Tradução: Rafael Zanolli. São Paulo: Novatec, 2015.
- ELMASRI, R.; NAVATHE, S. B. **Sistemas de Banco de Dados**. Tradução: Marília G. Pinheiro. 4. ed. São Paulo: Pearson, 2009.
- FERRARI, F. C. **Crie Banco de Dados em MySQL**. São Paulo: Digerati Books, 2007.
- GRAVES, Mark. **Projeto de Banco de Dados com XML**. São Paulo: Pearson, 2003.

HEUSER, C. A. **Projeto de Bancos de Dados**. Porto Alegre: Editora Bookman, 2009.

LAUDON, K.; LAUDON, J. **Sistemas de informações gerenciais**. Tradução: Célia Taniwaki. 11. ed. São Paulo: Pearson, 2014.

MACHADO, F. N. R. **Projeto e implementação de banco de dados**. 4. ed. São Paulo: Érica, 2020.

TAURION, C. **Big Data: velocidade, volume, variedade, veracidade, valor**. 1. Ed. Rio de Janeiro: Brasport, 2013.

Disciplina: Governança de dados

Ementa: Visão geral da gestão de dados, governança de dados, gestão da arquitetura de dados, desenvolvimento de dados, gestão de dados mestre e dados de referência.

Competências e Habilidades:

- Compreender o ciclo de vida do dado e o papel da governança
- Elaborar projetos usando princípios de governança de dados
- Esquematizar arquitetura de dados
- Aplicar princípios da gestão de dados

Conteúdo Programático 1: Introdução à governança de dados. Dado como ativo corporativo. Ciclo de vida do dado. Gestão de dados. O profissional de governança de dados.

Conteúdo Programático 2: Visão geral da gestão de dados. Missão, princípios e objetivos. Funções e atividades. Papéis e responsabilidades. Tecnologias.

Conteúdo Programático 3: Conceito de governança de dados. Conceitos e atividades. Detalhes de atividades da governança de dados. Desenvolvimento de dados. Estilos de modelagem de dados. Design de dados. Implementação de dados.

Conteúdo Programático 4: Arquitetura de dados. Gestão da arquitetura de dados. Conceitos e atividades. Princípios gerais. Gestão de dados mestre. Conceitos de dado mestre e dado de referência. Integração de dados. Gestão de mudanças.

Bibliografia Básica:

BRACKETT, Michael; EARLEY, Production Susan. **The DAMA Guide to The Data Management Body of Knowledge** (DAMA-DMBOK Guide), 2009.

BARBIERI, Carlos. **Governança de Dados:** Práticas, conceitos e novos caminhos. Alta Books, 2020.

RÊGO, Bergson Lopes. **Gestão e governança de dados:** promovendo dados como ativo de valor nas empresas. Brasport, 2013.

SHARDA, Ramesh; DELEN, Dursun; TURBAN, Efraim. **Business Intelligence e Análise de Dados para Gestão do Negócio.** Porto Alegre: Bookman, 2019.

Bibliografia Complementar:

HASHEM, I. A. T.; YAQOOB, I.; ANUAR, N.B.; MOKHTAR, S.; GANI, A.; KHAN, S.U. **The rise of “big data” on cloud computing: review and open research issues.** Information Systems 47 (2015), 98 - 115.

LOSHIN, David. **Master data management.** Morgan Kaufmann, 2010.

BAKSHI, Kapil. Considerations for big data: Architecture and approach. In: **2012 IEEE aerospace conference.** IEEE, 2012. p. 1-7.

Disciplina: Bancos de dados não relacionais (NoSQL).

Ementa: Conceituação de NoSql, prática e criação de bancos de dados não relacionais utilizando bancos comerciais de mercado como o MongoDB e o Hadoop. Estudo de formas de modelar bancos de dados não correlacionais.

Competências e Habilidades:

- Modelar banco de dados não correlacionais.

- Compreender comandos NoSQL.
- Verificar a utilização de segurança em banco de dados.
- Compreender integridade de dados.

Conteúdo Programático 1: Banco de Dados não-convencional: Introdução aos principais modelos de dados NoSQL;

Conteúdo Programático 2: NOSQL: Modelo de Sintaxe e dados JSON e XML.

Conteúdo Programático 3: Estrutura do MongoDB e Hadoop e seu uso prático.

Conteúdo Programático 4: Banco de Dados NoSQL: Modelo Orientado a documentos;

Conteúdo Programático 5: Banco de Dados NoSQL: Modelo Orientado a chave/valor;

Conteúdo Programático 6: Banco de Dados NoSQL: Modelo Orientado a família de colunas;

Conteúdo Programático 7: Banco de Dados NoSQL: Modelo Orientado a grafos;

Bibliografia Básica:

SADALAGE, P. J.; FOWLER, M. **NoSQL essencial**: um guia conciso para o mundo emergente da persistência poliglota. 1ed. São Paulo: Novatec, 2013.

SOUZA, M. B. **Desvendando o MongoDB**. Do Mongo Shell ao Java Driver. 1ed. São Paulo: Ciência Moderna, 2015.

SOUZA, M. B. **Introdução ao MONGODB**. São Paulo: Novatec, 2015

Bibliografia Complementar:

PANIZ, D. **NoSQL**: Como Armazenar os dados de uma aplicação moderna. São Paulo: Casa do Código, 2016.

TRICANO, A. **MongoDB**: NoSQL na prática. São Paulo: André Tricano, 2015.

PICHILIANI, C. A. **Introdução do MongoDB**: O ponto de partida para quem quer trabalhar com o banco de dados NoSQL orientado a documentos. São Paulo: Amazon, 2015

Disciplina: *Machine Learning*.

Ementa: Algoritmos de aprendizado iterativo a partir dos dados. Modelos preditivos. Redes neurais e conceitos de deep learning. Sistemas de recomendação.

Competências e Habilidades:

- Conceber modelos preditivos.
- Conhecer métodos de decisão e modelos.
- Aplicar inteligência artificial.

Conteúdo Programático 1: Introdução ao Machine Learning. Como funciona a aprendizagem de máquina.

Conteúdo Programático 2: Algoritmos de Machine Learninge modelos preditivos.

Conteúdo Programático 3: Regressão, K-Nearest Neighbours (KNN) e Naives Bayes.

Conteúdo Programático 4: Decision Tree, Random Forest e Método Ensemble.

Conteúdo Programático 5: Clusteing, Support Vector Machines e processamento em linguagem natural.

Conteúdo Programático 6: Redes neurais artificiais e introdução ao Deep Learning.

Conteúdo Programático 7: Sistemas de recomendação.

Bibliografia Básica:

FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. C. P. L. F. Inteligência artificial: uma abordagem de aprendizado de máquina. São Paulo: LTC Editora, 2011.

MITCHEL, M. B. Machine Learning. New York: McGraw-Hill Science/Engineering/Math, 1997. 432 p.

RASCHKA, Sebastian. Python Machine Learning. Birmingham: Packt Publishing, 2015.

FIGURELLI, R. Laboratórios de robôs: inteligência evolutiva para Machine Learning. Porto Alegre: Trajecta, 2017.

HOWARD, Mark. Machine Learning and Deep Learning: Essential User Guide to Learn and Understand Machine Learning and Deep Learning Effectively. Seattle: Amazon, 2017.

Disciplina: Projeto em ciência de dados com soluções para processamento paralelo e distribuído de dados.

Ementa: A disciplina traz o conceito de processamento paralelo e distribuído principalmente de consulta. Apresentar as arquiteturas existentes que trabalham atualmente com Sistemas Distribuídos.

Competências e Habilidades:

- Compreender arquitetura de processamento paralelo.
- Conhecer processamento distribuído.
- Aplicar conceitos do curso.

Conteúdo Programático 1: Introdução a Soluções para Processamento Paralelo e Distribuído de Dados. Gerenciamento de Clusters e Grids.

Conteúdo Programático 2: Monitoramento e depuração de programas paralelos. Sistemas de armazenamento de dados distribuídos.

Conteúdo Programático 3: Sistemas Peer-to-Peers para processamento Paralelo e Distribuído. Sistemas de Bancos de Dados Distribuídos principais características.

Conteúdo Programático 4: Construindo um Data Warehouse a partir de indicadores. Construindo ETL's para uso com o Data Warehouse criado.

Conteúdo Programático 5: Definindo um projeto de Big Data e as bases e arquivos a considerar. Tratando os arquivos de dados com Python.

Conteúdo Programático 6: Criando e tratando processos utilizando o R Programming. Criando e definindo Data Lakes a partir das bases disponibilizadas para tratativas.

Conteúdo Programático 7: Unificando e integrando os dados. Concluindo o projeto e exibindo os Dashboards em ferramentas OLAP

Bibliografia Básica:

KIMBALL, R. **The Data Warehouse Toolkit:** guia completo para modelagem dimensional. Rio de Janeiro: Campus, 2002.

BARI, A. **Integração de Dados na Prática.** Working with Big Data: Scaling Data Discovery. Seattle: Amazon Digital Services LLC, 2017.

GRUS, J. **Data Science do Zero:** Primeiras regras com o Python. Rio de Janeiro: Alta Books, 2016.

MAYER-SCHONBERGER, V. C. K. **BIG DATA:** Como Extrair Volume, Variedade, Velocidade e Valor da Avalanche de Informação Cotidiana. Rio de Janeiro: Elsevier Editora Ltda, 2013.

TANENBAUM, A. S. **Sistemas Distribuídos:** Princípios e Paradigmas. 2ed. São Paulo: PEARSON. 2015.

DANTAS, M. **Computação Distribuída de Alto Desempenho:** Redes, Clusters e Grids Computacionais. Seattle: Amazon, 2017.

COLOURIS, G.; DOLLIMORE, J.; KINDBERG, T. **Sistemas Distribuídos:** Conceitos e Projeto. 5ed. São Paulo: Bookman, 2013.

SADALAGE, P. J.; FOWLER, M. **NoSQL essencial**: um guia conciso para o mundo emergente da persistência poliglota. 1ed. São Paulo: Novatec, 2013.

MATTHES, E. **Curso intensivo de Python**: uma introdução prática e baseada em projetos à programação. São Paulo: Novatec Editora, 2017.

FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. C. P. L. F. **Inteligência artificial**: uma abordagem de aprendizado de máquina. São Paulo: LTC Editora, 2011.

COLOURIS, G.; DOLLIMORE, J.; KINDBERG, T. **Sistemas distribuídos**: conceitos e projeto. 5ed. São Paulo: Bookman, 2013.

BAEZA-YATES, R.; RIBEIRO N. B. **Recuperação de informação**: conceitos e tecnologia das máquinas de busca. 2ed. São Paulo: Bookman, 2013.

CAMARGO, A.; AMARAL, E.; BASTOS, R. **Bancos de dados distribuídos**: uma solução aplicada: Implemente e gerencie um SGBDD! São Paulo: Novas Edições Acadêmicas, 2017.

VALDURIEZ, P.; OZSU, M. T. **Princípios de sistemas de bancos de dados distribuídos**. Rio de Janeiro: Elsevier, 2001.

DANTAS, M. **Redes de comunicação e computadores** – abordagem quantitativa. São Paulo: Amazon.

Disciplina: Administração de bancos de dados.

Ementa: Processamento de transações. Controle de concorrências. Técnicas de backup e recuperação de banco de dados. Ferramentas de gerenciamento para SQL Server e Oracle. Administração de usuários. Segurança de banco de dados. Logs. Sistemas de banco de dados distribuídos.

Competências e Habilidades:

- Criar projetos de banco de dados
- Operar bancos de dados relacionados como administrador de sistemas

- Praticar conceitos de banco de dados em SGBDs de mercado
- Gerir bancos de dados e seus usuários

Conteúdo Programático 1: Sistemas gerenciadores de bancos de dados. SQL Server. Ferramentas de gerenciamento. SQL Server 2019 Express. Configuration Manager, Management Studio e SQLCMD. Arquitetura do banco de dados Oracle. Oracle 10g Express Edition. Ferramentas sqlplus e SQL Developer.

Conteúdo Programático 2: Processamento de transações. Armazenamento e leitura de logs de transações. Estatísticas. Aplicações em SQL Server. Aplicações em Oracle. Gerenciamento de transações. Controle de concorrência. Falhas e propriedades ACID. Transações, locks, blocks e deadlocks.

Conteúdo Programático 3: Técnicas de backup e recuperação de banco de dados. Gerenciamento do espaço em disco. Aplicações em SQL Server. Aplicações em Oracle.

Conteúdo Programático 4: Administração de usuários. Logins, usuários e permissões. Controle de acesso. Segurança de banco de dados. SQL Injection. Criptografia. Comandos em linguagem SQL. Aplicações em SQL Server. Aplicações em Oracle.

Bibliografia Básica:

ELMASRI, Ramez; NAVATHE, Shamkant B. **Sistemas de banco de dados**. 4ed. São Paulo: Pearson, 2009.

LEBLANC, Patrick. **Microsoft SQL Server 2012: Passo a Passo**. Porto Alegre: Bookman, 2014.

WATSON, John. **OCA Oracle Database 11g: Administração I**. Porto Alegre: Bookman, 2014.

Bibliografia Complementar:

DATE, Christopher J. **Introdução a sistemas de bancos de dados**. 8. ed. Rio de Janeiro: Elsevier, 2004.

DE SOUZA, E. P. **Migração de banco de dados Oracle para PostgreSQL**. **Embrapa Informática Agropecuária**. Documentos, 2007.

DOS SANTOS, Alex Lima et al. FUNCIONAMENTO DO BANCO DE DADOS RELACIONAL SQL SERVER 2000 ENTERPRISE. **Revista de Informática Aplicada**, v. 3, n. 1, 2007.

Disciplina: Gerenciamento do desempenho do banco de dados (tuning).

Ementa: Análise e aperfeiçoamento das regras de negócio, visando o máximo desempenho das aplicações. Exame do desempenho de consultas que utilizam filtros, através do funcionamento do otimizador e seus planos de execução. Averiguação e otimização das estruturas em memória. Utilização de estruturas físicas diferenciadas para casos especiais.

Competências e Habilidades:

- Conhecer, analisar e aprimorar as regras e recursos dos sistemas de banco de dados em operação;
- Saber observar e modificar consultas que utilizam filtros, a fim de melhorar o desempenho;
- Conhecer e saber utilizar planos de execução de um otimizador;
- Estimar, inferir, e analisar estatisticamente o desempenho das consultas que utilizam índices.
- Conhecer e adequar as estruturas em memória, além de elaborar estruturas físicas diferenciadas para casos especiais.

Conteúdo Programático 1: Entendendo a estratégia de ajuste de performance como um método. Analisando e aperfeiçoando as regras de negócio, adequando o uso das aplicações, o projeto dos dados e os recursos do banco de dados.

Conteúdo Programático 2: Examinando o desempenho de consultas que utilizam filtros. Análise do otimizador e os planos de execução. Medindo o impacto do desempenho das consultas com índices, através do otimizador.

Conteúdo Programático 3: Conhecendo e adequando o desempenho das estruturas em memória.

Conteúdo Programático 4: Utilizando estruturas físicas diferenciadas para conjuntos especiais de dados.

Bibliografia Básica:

ELMASRI, Ramez; NAVATHE, Shamkant B. **Sistemas de banco de dados**. 7. ed. São Paulo: Pearson Education, 2019.

DATE, Christopher J. **Introdução a sistemas de bancos de dados**. 8. ed. Rio de Janeiro: Elsevier, 2004.

SILBERSCHATZ, Abraham; KORTH, Henry F.; SUDARSHAN, S. **Sistema de banco de dados**. 6. ed. Rio de Janeiro: Elsevier, 2012

Bibliografia Complementar (inserir 3 bibliografias atuais):

RAMAKRISHNAN, Raghu. **Sistemas de gerenciamento de banco de dados**. 3. ed. Nova Iorque: McGraw Hill, 2011.

FLORENTINO, Plínio T. **Gerenciamento e desenvolvimento em banco de dados**. Londrina: Editora e Distribuidora Educacional, 2018.

HEUSER, Carlos A. **Banco de dados relacional: conceitos, SQL e administração**. Porto Alegre: Clube de Autores, 2019.

13. Infraestrutura Física e Pedagógica

O aluno encontrará todo o conteúdo do curso e assistirá às aulas gravadas no ambiente virtual. Para assistir às aulas é fundamental que as especificações abaixo sejam obedecidas, possibilitando, assim, uma recepção de maior qualidade dos vídeos.

Hardware:

- Processador Intel Core 2 Duo ou superior.
- 2Gb de Memória RAM.

- Placa de vídeo com resolução 1024x768, qualidade de cor 32 bits e compatível com Microsoft DirectShow.
- Microsoft DirectX 9.0c ou posterior.

Software:

- Navegador: Firefox, Google Chrome, Internet Explorer (sempre atualizado).
- Sistema Operacional: Windows XP ou posterior.
- Adobe Flash Player (atualizado).
- *Plugin* de vídeos SilverLigth (atualizado)

Rede:

- Conexão com a Internet banda larga de no mínimo 2 MB.
- Em caso de acesso em ambientes corporativos além da velocidade, é necessário verificar as condições de segurança de rede de sua empresa e se certificar que o site não estará bloqueado.

Adicionalmente, é prevista a utilização da biblioteca virtual para consultas bibliográficas e pesquisa de assuntos referentes às disciplinas ministradas.

