

<b>Plano de Desenvolvimento da Disciplina</b>	
MC936 - Ciência e Visualização de Dados em Saúde Instituto de Computação Universidade Estadual de Campinas	2024/1  Professor: André Santanchè

## Horários

Terça-feira: 8:00 às 10:00

Quinta-feira: 8:00 às 10:00

## Ementa

Curso com base interdisciplinar que acomoda diferentes trilhas de aprendizado, possibilitando a participação e interação de alunos de ciências exatas (e.g., computação, engenharias, matemática e estatística), saúde (e.g., medicina e farmácia) e biologia. Apresenta um panorama de todas as etapas do processo de descoberta de conhecimento, com aplicações no contexto da saúde: (1) estratégias de seleção de dados guiadas por perguntas de pesquisa; (2) técnicas de pré-processamento; (3) etapas típicas de transformação; (4) análise e mineração de dados; (5) interpretação e avaliação por meio de técnicas de visualização de dados. O curso adota uma abordagem prática com apresentação de ferramentas e seu uso na análise de dados, bem como atividades e um projeto em que interagem alunos de ciências exatas e saúde/biologia. Durante todo o curso, são tratadas especificidades de domínio da ciência de dados no contexto da saúde.

## Programa

A análise computacional de dados em saúde tem crescido em relevância e impacto não apenas pelo desenvolvimento de estratégias e algoritmos, como também pela crescente quantidade de dados disponíveis. Este curso tem base interdisciplinar e envolve professoras e professores de computação, engenharia elétrica, biologia e medicina. Ela acomodará diferentes trilhas de aprendizado, possibilitando a participação e interação de alunos de ciências exatas (e.g., computação, engenharias, matemática e estatística), biologia e saúde (e.g., medicina, enfermagem e farmácia). A disciplina terá a participação de professores colaboradores, que fazem parte do staff.

Neste ano, a disciplina dará o enfoque na saúde sobre a perspectiva de Biologia de Sistemas e sua representação/análise sob o enfoque de grafos/redes.

A disciplina alinhará alunos com conceitos fundamentais de biologia de sistemas, como:

- transcrição
- tradução
- regulação transcricional
- regulação pós-transcricional
- gene regulatory networks (GRN)

Do ponto de vista de Ciência de Dados a disciplina abordará:

- estratégias de seleção de dados guiadas por perguntas de pesquisa
- técnicas de obtenção, pré-processamento e transformação de dados
- estratégias de exploração e análise de dados da área da saúde

Em relação à análise de dados com enfoque em ciência de redes, serão abordados os tópicos:

- representando fenômenos em grafos/redes
- redes de conhecimento e ontologias
- redes complexas

- ciência de redes
- grafos de conhecimento e aprendizagem de máquina

A disciplina é guiada por problemas em saúde trazidos por professores da biologia e medicina. Nesta edição, já estão sendo preparados problemas no domínio de câncer e hematologia, mas poderão sugerir problemas em outros domínios.

## Critérios de Avaliação

---

O curso terá as seguintes avaliações e respectivos períodos:

Legenda	Descrição
PC	Participação no Curso.
Lab	Exercícios práticos feitos no laboratório.
Prj	Projeto da disciplina desenvolvido ao longo do semestre.

A especificação do projeto e dos laboratórios serão entregues em documentos específicos. O projeto terá datas de entrega parciais que serão definidas em sala durante o curso.

O cálculo da média é calculada como:

$$\text{média} = (\text{PC} * 1 + \text{Lab} * 5 + \text{Prj} * 4) / 10$$

### Exame Final

Estarão dispensados do exame apenas os alunos com média  $\geq 5$

Para estar habilitado a realizar o exame o aluno deve ter média mínima: média  $\geq 2,5$

Data de realização: 11/07/2024 às 8h

Neste caso, o cálculo da média para alunos que precisam do exame:

$$\text{média}_{\text{final}} = (\text{média} + \text{nota}_{\text{exame}}) / 2$$

## Bibliografia

---

Alberts, B., Heald, R., Johnson, A. (Alexander D. ), Morgan, D. O., Raff, M. C., Roberts, K. (Keith), & Walter, P. (2022). Molecular Biology of the Cell. W.W. Norton & Company, 661.

Consoli, Sergio, Diego Reforgiato Recupero, and Milan Petkovic (2019). Data Science for Healthcare. Springer International Publishing.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2000). An introduction to Statistical Learning. Springer.

Klipp, L., Wierling, K., Lehrach & Herwig. 2009. Systems Biology: A Textbook. VCH/John Wiley.

Munzner, Tamara (2014). Visualization Analysis and Design. CRC Press.

Palsson, B. 2006. Systems Biology: Properties of Reconstructed Networks. Cambridge University Press.

Piatetski, Gregory, and William Frawley (1991). Knowledge discovery in databases. MIT press.

Shortliffe, E. H., & Cimino, J. J. (Eds.). (2014). Biomedical Informatics - Computer Applications in Health Care and Biomedicine. London: Springer London.