Summer Internship 2022 IC/UNICAMP

AI R&D Team

- Os objetivos, áreas de interesse e cronograma de cada vaga foram separados por projeto (descrição logo abaixo).

- **Remuneração proposta:** R$ 2000,00 à R$3000,00

Vantagens adicionais:
Vale Refeição, Plano de Saúde, Plano odontológico, Transporte, Estacionamento, Portal de vendas produtos Samsung.

- **Descrição do programa:** The candidate will work with our team of researchers, data scientists and health domain specialists in order to design and deliver innovative health features to Samsung devices and services. The candidate will work on projects that affect the life of millions of people worldwide, helping them to improve their health and wellbeing.

- **Resumo das vagas:** total 13 vagas

Project	Area	Grad level	# positions
1 - MIA	Data science	Master or Doctorate	1
1 - MIA	Research	Master or Doctorate	1
1 - MIA	Development	Master or Doctorate	1
2 - CRF	Research + Data Science	Master or Doctorate	2
2 - CRF	Data Science	Master or Doctorate	1
3 - BP	Development	Master or Doctorate	1
4 - PeHR	Research	Doctorate	1

Project 1 - MIA

Total number of positions: 3

Positions 1, 2, and 3:

Proposal: AI researcher intern (1), Data scientist intern (2), AI development intern (3)

Objectives: The candidate will be part of our awesome team of engineers, data scientists, and researchers to build cutting edge technologies and implementing the latest research into our products impacting the life of millions of people worldwide.

Areas of interests: The candidate will have the opportunity to work in one or more out of three areas of specialization:

- Model embedding
- Wearable sensors study
- Machine learning modeling

Grad level: Master or doctorate

Schedule:

Week 1:

- Presenting the company, the team, and the project
- Presenting the problems and goals related to the project
- Intern software engineering training
- Creating and preparing the work environment

Week 2:

- Introducing the main concepts related to the project
- Introducing the tools and frameworks employed by the team
- Introducing the code and pipelines developed by the team
- Introducing the data to be used

Week 3:

- Literature review
- Defining the problem and goals

Weeks 4 and 5:

- (1) Researcher - Machine learning modeling
    - Bayesian modeling
    - Neural Architecture Search (NAS)
- (2) Data science - Wearable sensors study
    - Feature engineering
    - Data augmentation
- (3) Developer - Model embedding
    - Model compression
    - Transpiling (source-to-source)

Weeks 6:

- Study Presentation
- Designing and coding experiments
- Analyzing and improving the proposed approaches

Weeks 7, 8 & 9:

- Reviewing the goals and proposed solutions
- Exploring other ways to solve the problem
- Comparing results with other known approaches
- Integrating proposed solution into framework

Weeks 10, 11 & 12:

- Writing a technical report /scientific paper
- Writing documentation
- Presenting the final results to the team

Project 2 - CRF

Total number of positions: 2

Positions 1 and 2:

Proposal: R&D AI intern

Objectives: The candidate will work on a health related solution proposed to Samsung's devices that combines data collection, processing and evaluation based on machine learning techniques. The candidate will face a problem related to analyzing and processing the gathered health data, aiming to filter it to enhance the design of a solution built of machine learning models. We expect the candidate to improve his/her skills and knowledge in the areas of data analysis and processing, as well as the evaluation of machine learning methods based on literature research, proposing new features (or composition of features) to feed machine learning models to enhance the solution currently in development inside the team.

Areas of interest:

- Machine learning
- Data science
- Health science

Grad. level: Master or Doctorate

Schedule:

Week 1:

- Presenting the company, the team, and the project
- Presenting the problems and goals related to the project
- Presenting the tools and frameworks employed by the team

- Creating and preparing the work environment

Week 2 and 3:

- Introducing the main concepts related to the project
- Introducing the code and pipelines developed by the team
- Introducing the machine learning algorithms developed by the team
- Introducing the data used to train the models

Weeks 4, 5 and 6:

- Defining the problem and goals
- Literature review concerning the defined problem and goals
- Getting in touch with the data, its nuances and configuration
- Data cleaning and preparation do deal with the tasks

Weeks 7, 8, and 9:

- Designing and code the models using our pipelines and frameworks
- Training, evaluating, and updating the developed models
- Comparing results with other known models
- Analyzing and improving the proposed model
- Reviewing the goals and proposed solutions

Weeks 10, 11 and 12:

- Analysis and review of the results obtained so far
- Tests and code review
- Writing a scientific paper/technical report
- Writing documentation
- Presenting the final results to the team

Position 3:

Proposal: Data scientist intern

Objectives: The candidate will work on a health related solution proposed to Samsung's devices that combines data collection, processing, analysis and evaluation based on machine learning techniques. The candidate will face a problem related to data visualization. We expect the candidate to improve

his/her skills and knowledge in the areas of development, data processing, aggregation, analysis and data storytelling.

Areas of interest:

- Data science
- Data visualization

Grad. level: Master or Doctorate

Schedule:

Week 1:

- Presenting the company, the team, and the project
- Presenting the problems and goals related to the project
- Presenting the tools and frameworks employed by the team
- Creating and preparing the work environment

Week 2 and 3:

- Introducing the main concepts related to the project
- Introducing the data used by the team
- Introducing the data processing pipeline developed by the team

Weeks 4, 5 and 6:

- Defining the problem and goals
- Getting in touch with the data, its nuances and configuration
- Data cleaning and preparation to deal with the tasks
- Framework and tools necessary to development

Weeks 7, 8, and 9:

- Designing and code the models using our pipelines and frameworks

Weeks 10, 11 and 12:

- Analysis and review of the results obtained so far
- Tests and code review
- Writing a technical report
- Writing documentation
- Presenting the final results to the team

Project 3 - BP

Total number of positions: 1

Proposal: AI development intern

Objectives: The candidate will be part of our awesome team of engineers, data scientists, and researchers to build innovative technologies and implementing the latest research into our products influencing the life of millions of people worldwide.

Areas of interests:

- Machine learning
- Development (Python and C)

The candidate will have the opportunity to work in model embedding, where the effort is the reduction of footprint. Python and C programming.

Graduate level: Master or doctorate

Schedule:

Week 1:

- Presenting the company, the team, and the project
- Presenting the problems and goals related to the project
- Intern software engineering training
- Creating and preparing the work environment

Week 2:

- Introducing the main concepts related to the project
- Introducing the tools and frameworks employed by the team

- Introducing the code and pipelines developed by the team
- Introducing the data to be used

Week 3:

- Literature review
- Defining the problem and goals

Weeks 4 & 5:

- Model embedding
    - Model compression
    - Transpiling (source-to-source)

Weeks 6:

- Study Presentation
- Designing and coding experiments
- Analyzing and improving the proposed approaches

Weeks 7, 8 & 9:

- Reviewing the goals and proposed solutions
- Exploring other ways to solve the problem
- Comparing results with other known approaches
- Integrating proposed solution into framework

Weeks 10, 11 & 12:

- Writing a technical report /scientific paper
- Writing documentation
- Presenting the final results to the team

Project 4 - PeHR

Total number of positions: 1

Position 1:

Proposal: R&D AI intern

Objectives: Acquiring on-hands experience in deploying AI solutions in low-memory wearable devices and converting machine learning models between high level and low level languages. It is expected that the candidate develop skills and knowledge in the direct implementation and testing of machine learning methods, as well as have the opportunity to study the different steps related to AI development and deployment such as Python -> C transpilation.

Areas of interest:

- Machine learning
- Data science
- Health science

Grad level: Doctorate

Schedule:

Week 1:

- Presentation of the company, the team, and the project
- Presenting the problems and goals related to the project
- Presenting the tools and frameworks employed by the team
- Creating and preparing the work environment

Week 2:

- Presentation of the two main efforts (sub-projects) in the project
- Presentation of the Python -> C transpilation efforts in project
- Presentation of the efforts related to the unification of the two sub-projects' APIs (Python)
- Presentation of the efforts related to the unification of the two sub-projects' APIs (C)

Week 3:

- Implementation of a Python -> C transpilator for machine learning methods used in project
    - Linear regression methods (Scikit Learn)
    - Tests on the two sub-projects' libraries and code review

Weeks 4 and 5:

- Implementation of a Python -> C transpilator for machine learning methods used in the project
    - MultiLayer Perceptron-based models (Pytorch, Tensorflow, Scikit Learn)
    - Tests on the two sub-projects' libraries and code review

Weeks 6 and 7:

- Implementation of a Python -> C transpilator for machine learning methods used in the project
    - Tree and ensemble models (random forest, gradient boosted trees, XGBRegressor) (Scikit Learn, XGBoost)
    - Tests on two sub-projects' libraries and code review

Weeks 8 and 9:

- Unification of developed Python -> C transpilators for ML methods into a single API
- Tests on the two sub-projects' libraries and code review

Week 10, 11 and 12:

- Unification of the two sub-projects' libraries into a single API
- Tests and code review
- Presenting the final results to the team