



FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO

PLANO DE ATIVIDADES INDIVIDUAL PARA BOLSA DE TREINAMENTO TÉCNICO E PARTICIPAÇÃO EM CURSO (fapesp.br/tt)

Escolha o Nível Pretendido (somente uma opção por formulário)

<input type="checkbox"/> TT-1	<input type="checkbox"/> TT-2	<input checked="" type="checkbox"/> TT-3	<input type="checkbox"/> TT-4	<input type="checkbox"/> TT-4A	<input type="checkbox"/> TT-5
<input type="checkbox"/> PC-2		<input type="checkbox"/> PC-3		<input type="checkbox"/> PC-4	

Duração da bolsa (somente TT): 9(meses)	Dedicação (somente TT):40(horas/semana)
---	---

1) Dados do Projeto a que se vincula a bolsa:

Título: Modelo Inteligente de Gestão de Varejo
No. Processo:
Pesquisador Responsável: Douglas Diniz Landim
Instituição/Empresa: ATIDO

2) TÍTULO DO PLANO DE PESQUISA a ser desenvolvido pelo(a) bolsista:

Bolsa de Treinamento Técnico TT-3 para tratamento dos dados do projeto e codificação dos modelos de aprendizado de aprendizado de máquina desenvolvidos em pesquisa.
--

3) RESUMO do trabalho a ser desenvolvido:

<p>O bolsista se dedicará a atividade principal de operação das bibliotecas e tecnologias utilizadas pelo projeto, sendo todas da linguagem Python: TensorFlow com integração com o Keras, Scikit Learn e Pandas.</p> <p>Linguagem SQL com servidor MYSQL.</p> <p>Plataforma Jupyter Notebook</p> <p>Plataforma GIT.</p> <p>Receberá treinamento durante o projeto para a correta utilização das plataformas, e para todas as suas atividades de codificação receberá os modelos já determinados e projetados na atividade científica para a codificação, bem como o roteiro experimental de suas atividades.</p>

4) OBJETIVOS pretendidos:

<p>Os objetivos do bolsista em termos do desafio proposto é se familiarizar com as ferramentas do projeto para tratar os dados originados na pesquisa de campo e atividade de business intelligence, em formato que possa ser lido e processado pelas ferramentas de aprendizado de máquina.</p> <p>Receberá desenhos e modelos teóricos de topologias de redes neurais, com parâmetros determinados, com o objetivo de transformar os modelos teóricos fornecidos em código e experimentos, para tal receberá também um roteiro de atividades experimentais, treino e avaliação destes modelos de aprendizado de máquina.</p> <p>Também ficará responsável pela atualização e manutenção do repositório github do projeto, e documentação e relato de suas atividades.</p>

5) PLANO DE TRABALHO (inclua a Metodologia, as atividades a serem desempenhadas pelo bolsista e a Tabela com Cronograma de Execução das atividades).

Este plano de trabalho será conduzido por métodos da ATIDO adquiridos com sua experiência atividades de T.I que incorpora desenvolvimento ágil em seus processos e pelo estudo e aplicação do código das boas práticas científicas FAPESP e do estudo e aplicação da metodologia FASCATI, realizando treinamento teórico e experimental em ciclos, visando minimizar as lacunas técnicas e gerenciais usualmente encontradas na interação entre projeto teórico e desenvolvimento prático, de forma incremental.

Itens entregáveis ao longo de seu trabalho:

Tratamento e armazenamento dos dados coletados em formato computacional acessível para utilização nos modelos de aprendizado de máquina, Separação dos dados em conjuntos de treino, validação e teste ,em banco de dados MYSQL. Formato de Arquivo SQL, prazo 2/11/2020.

Adaptação dos dados dos conjuntos de dados, armazenados em MYSQL em formato de entrada para o modelo pesquisado. Codificação e execução de treino, com reajuste empírico, dos modelos específicos de rede neural artificial, para cada negócio de varejo pesquisado, com a adaptação do número de neurônios da camada de saída correspondente ao número de produtos distintos de cada negócio. Formato de Código python prazo 28/12/2020.

Codificação e treino do modelo global de rede neural artificial, em cópias distintas para cada negócio varejista, com única modificação na camada de saída baseada no número de produtos de cada negócio, treino e execução de cada cópia. Formato de Código python prazo 19/1/2020.

Adaptação dos conjuntos de dados dos negócios distintos, para a entrada da rede recorrente, e execução do treino da rede recorrente em formato MYSQL. prazo 5/3/2021,

Codificação e atualização no repositório das redes recorrentes, em formato de Código Python, prazo 22/04/2021

Cronograma Completo:

BOLSISTA	TT3 (ATIVIDADES DE PROGRAMAÇÃO)	COM OU PARA QUEM	ONDE	%	INÍCIO
TÉRMINO					
0.1 - Reunião de abertura: Videoconferência com os bolsistas e parceiro do projeto, com apresentação do plano de pesquisa e normas	TODOS	HANGOUTS	0%	seg, 3/8/2020	seg, 3/8/2020
0.1 - Estudo das Normas Pipe	PE	REMOTO	0%	seg, 3/8/2020	sex, 7/8/2020
0.1 - Estudo das boas práticas científicas	PE	REMOTO		sex, 7/8/2020	qua, 12/8/2020
0.1 - Estudo da metodologia frascati	PE	REMOTO		qui, 13/8/2020	qui, 20/8/2020
0.5 Integração de equipe e início de atividades na sede	TT3/TT4	SEDE	0%	sex, 21/8/2020	sex, 21/8/2020
3.2 - Estudo e treinamento da documentação das ferramentas Git, Python, SciKit Learning, Keras, Pandas e MYSQL	PE	SEDE	0%	seg, 24/8/2020	dom, 4/10/2020
3.4 - Tratamento e armazenamento dos dados coletados em MYSQL	PE	SEDE	0%	seg, 5/10/2020	seg, 19/10/2020
4.2 - Separação dos conjuntos de treino, validação e teste, por empresa pesquisada	PE/TT4	SEDE	0%	ter, 20/10/2020	dom, 25/10/2020
3.5 - Revisão e entrega do banco MYSQL FAPESP	SEDE	0%	seg, 26/10/2020	dom, 1/11/2020	
0.6 - Reunião e integração de equipe sobre o 2o trimestre	TODOS	SEDE	0%	seg, 2/11/2020	ter, 3/11/2020
5.2 - Revisão e treinamento das ferramentas do Keras/Scikit Learning	PE	SEDE	0%	qua, 4/11/2020	qua, 18/11/2020
6.1 - Adaptação da entrada do conjunto de dados para os modelos do item 5	PE/TT4	SEDE	0%	qui, 19/11/2020	qui, 26/11/2020
6.2 - Codificação dos modelos específicos, adaptação da camada de saída	PE/TT4	SEDE	0%	sex, 27/11/2020	qui, 3/12/2020
6.3 - Treino e Reajuste Empírico dos Modelos Específicos	PE/TT4	SEDE	0%	sex, 4/12/2020	sex, 18/12/2020
6.4 - Revisão dos códigos e atualização dos repositórios	PE	SEDE	0%	seg, 21/12/2020	qua, 23/12/2020
8.1 - Codificação do modelo global, treino e execução.	PE/TT4	SEDE	0%	sex, 8/1/2021	qui, 14/1/2021
8.2 - Revisão dos códigos do modelo global e atualização dos repositórios	PE	SEDE	0%	qui, 14/1/2021	sex, 15/1/2021
10.1 - Treinamento técnico de codificação de redes LSTM	PE	SEDE	0%	seg, 18/1/2021	sex, 31/1/2020
0.8 - Reunião e integração de equipe sobre o 3o trimestre	TODOS	SEDE	0%	seg, 1/2/2021	ter, 2/2/2021
10.1 - Adaptação dos conjuntos de dados dos negócios	PE/TT4	SEDE	0%	qua, 3/2/2021	sex, 5/3/2021
11.2 - Codificação e atualização no repositório das redes recorrentes	PE/TT4	SEDE	0%	seg, 8/3/2021	qua, 7/4/2021
11.4 - Treino e aplicação de métricas avaliativas da rede recorrente	PE/TT4	SEDE	0%	qui, 8/4/2021	qui, 22/4/2021
12.4 - Elaboração do relatório científico parcial, de todos os experimentos	PE/TT3	SEDE	0%	qui, 8/4/2021	qua, 21/4/2021
13.1 - Elaboração do relatório científico final da fase 1	PE/TT3	SEDE	0%	sex, 23/4/2021	qui, 29/4/2021
0.9 - Reunião de equipe e encerramento das atividades da fase 1	TODOS	SEDE	0%	sex, 30/4/2021	sex, 30/4/2021

6) Justifique a escolha do Nível para a bolsa (TT 1,2,3,4,4A,5 / PC 2,3,4).

Para as atividades do bolsista foi escolhido o nível TT3, para aluno graduado do nível superior na área de T.I inerente às suas das atividades de apoio ao projeto de pesquisa, sem a necessidade de experiência ou especialização técnica em aprendizado de máquina, pois receberá ao longo do projeto atividades de treinamento técnico de utilização de suas ferramentas, bem como projetos de modelos de aprendizado de máquina e roteiros de treinamento e avaliação destes modelos.

7) Justifique o Plano de Atividades em termos dos objetivos do Programa de Bolsas de Capacitação Técnica (apoio técnico ao projeto de pesquisa e treinamento técnico do bolsista).

As atividades do cronograma do bolsista TT3 correspondem aos objetivos do programa de capacitação técnica, dando treinamento técnico e capacitação técnica das principais ferramentas de aprendizado de máquina requisitadas pelo mercado.

No fim do projeto de pesquisa o bolsista terá experiência para elaborar uma análise de dados de mercado desde sua concepção até a entrega de resultados, adquirindo uma capacitação técnica de um setor muito promissor no ramo de T.I. Garantir apoio técnico aos projetos de pesquisa apoiados pela FAPESP.

É importante ressaltar que o desafio científico não está no desenvolvimento de atividades realizado pelo bolsista tt3 que receberá treinamento e suporte da PE, mas sim na pesquisa científica e na modelagem teórica dos modelos de aprendizado de máquina e que exigem conhecimentos acadêmicos especializados.

8) Informações adicionais

8.1) Observa-se que as bolsas TT4, TT4A e TT5 destinam-se a profissionais de TI (Tecnologia da Informação), para desenvolver atividades de TI (p.ex., desenvolvimento de software, integração de sistemas de software, projeto e análise de algoritmos, modelagem e implementação de bancos de dados complexos, arquitetura de software, projeto e implementação de sistemas embarcados e microcontrolados) junto ao projeto de pesquisa. As atividades de TI devem estar precisamente descritas. O tempo de experiência do candidato é contado após a conclusão da graduação.

8.2) Caso sejam solicitadas várias bolsas TT, os planos devem ser individuais, sem sobreposição de tarefas. Caso haja várias bolsas, é recomendada também a inclusão de um documento de apresentação descrevendo sucintamente como as bolsas se inter-relacionam e se complementam de forma a ficar claro a sua necessidade para a execução da pesquisa



FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO

PLANO DE ATIVIDADES INDIVIDUAL PARA BOLSA DE TREINAMENTO TÉCNICO E PARTICIPAÇÃO EM CURSO (fapesp.br/tt)

Escolha o Nível Pretendido (somente uma opção por formulário)

<input type="checkbox"/> TT-1	<input type="checkbox"/> TT-2	<input type="checkbox"/> TT-3	<input checked="" type="checkbox"/> TT-4	<input type="checkbox"/> TT-4A	<input type="checkbox"/> TT-5
<input type="checkbox"/> PC-2		<input type="checkbox"/> PC-3		<input type="checkbox"/> PC-4	

Duração da bolsa (somente TT): 9(meses)	Dedicação (somente TT):40(horas/semana)
---	---

1) Dados do Projeto a que se vincula a bolsa:

Título: Modelo Inteligente de Gestão de Varejo
No. Processo:
Pesquisador Responsável: Douglas Diniz Landim
Instituição/Empresa: ATIDO

2) TÍTULO DO PLANO DE PESQUISA a ser desenvolvido pelo(a) bolsista:

Bolsa de Treinamento Técnico TT-4 para pesquisa científica e desenvolvimento de modelos de aprendizado de máquina para previsão de demanda de consumo com dados exógenos.

3) RESUMO do trabalho a ser desenvolvido:

Bolsista de atividades científicas específicas de aprendizado de máquina, essenciais ao projeto, como revisões e expansões bibliográficas, modelagem de redes neurais, recorrentes, metodologias científicas, análise de dados, entre outras atividades essenciais em serem realizadas de forma simultânea ao trabalho técnico de codificação de tratamento dos dados, sendo necessário a cooperação e trabalho simultâneo entre equipe técnica e científica.

4) OBJETIVOS pretendidos:

<p>Os objetivos do bolsista TT4 em termos do desafio proposto é realizar trabalho científico de pesquisa e modelagem teórica constante ao longo do projeto, para identificar modelos e metodologias de aprendizado de máquina que atendam os objetivos do projeto.</p> <p>O bolsista TT4, irá entregar modelos teóricos, projetados e desenhados, ao desenvolvimento técnico que irá codificar e treinar e avaliar os modelos de acordo com as especificações e roteiros experimentais determinados pelo bolsista, através de sua pesquisa científica.</p> <p>Também poderá solicitar da equipe técnica tratamento de conjunto de dados para análises</p> <p>Ao longo de seu trabalho, irá produzir documentos e relatórios científicos, bem como acompanhar e orientar a equipe técnica</p>
--

5) PLANO DE TRABALHO (inclua a Metodologia, as atividades a serem desempenhadas pelo bolsista e a Tabela com Cronograma de Execução das atividades).

Este plano de trabalho será conduzido por métodos da ATIDO adquiridos com sua experiência atividades de T.I que incorpora desenvolvimento ágil em seus processos e pelo estudo e aplicação do código das boas práticas científicas FAPESP e do estudo e aplicação da metodologia FASCATI, realizando treinamento teórico e experimental em ciclos, visando minimizar as lacunas técnicas e gerenciais usualmente encontradas na interação entre projeto teórico e desenvolvimento prático, de forma incremental.

Itens entregáveis ao longo de seu trabalho:

Relatório de identificadores e dados exógenos de mercado, que possam ser processados pelos modelos de aprendizado, em trabalho conjunto com a equipe de business intelligence, prazo 20/10/2020.

Pesquisa e desenvolvimento de análise exploratória dos dados coletados identificação e aplicação de métodos estocásticos de previsão de demanda. Formato de Relatório, prazo 2/11/2020.

Discussão e análise de resultados entre modelo global e o específico de cada negócio. Formato de Relatório Científico, prazo 31/1/2021

Estudo e modelagem da topologia do modelo global de rede neural artificial com base na topologia dos modelos específicos. Formato de Relatório, prazo 12/1/2021

Pesquisa e desenvolvimento de modelo de aprendizado de máquina com rede recorrente LSTM. Formato de Relatório, prazo 5/3/2021

Elaboração da metodologia de treino e avaliação da rede recorrente. Formato de Relatório, prazo 7/4/2021

Relatório comparativo de todos os modelos estudados, prazo 21/4/2021

Cronograma Completo:

BOLSISTA TT4 (ATIVIDADES CIENTÍFICAS) COM OU PARA QUEM ONDE PROGRESSO INÍCIO
TÉRMINO

0.1 - Reunião de abertura: Videoconferência com os bolsistas e parceiro do projeto, com apresentação do plano de pesquisa e normas	TODOS	HANGOUTS	0%	seg, 3/8/2020	seg, 3/8/2020
0.1 - Estudo das Normas Pipe	PE	REMOTO	0%	seg, 3/8/2020	sex, 7/8/2020
0.1 - Estudo das boas práticas científicas	PE	REMOTO	0%	sex, 7/8/2020	qua, 12/8/2020
0.1 - Estudo da metodologia frascati	PE	REMOTO	0%	qui, 13/8/2020	qui, 20/8/2020
0.5 Integração de equipe e início de atividades na sede	TT3/TT4	SEDE	0%	sex, 21/8/2020	sex, 21/8/2020
4.1 - Estudo e expansão bibliográfica de análise exploratória de dados coletados, elaboração de metodologia de aplicação de métodos estocásticos de previsão de demanda.	PE/PBI	SEDE	0%	seg, 24/8/2020	qua, 23/9/2020
3.3 - Estudo, expansão bibliográfica e elaboração de metodologia da separação dos dados em conjunto de treino, validação e teste.	PE/PBI	SEDE	0%	qui, 24/9/2020	dom, 4/10/2020
4.3 - Conclusão do documento científico de separação do conjunto de dados, e de métodos estocásticos de previsão de demanda	PE	SEDE	0%	seg, 5/10/2020	dom, 1/11/2020
0.6 - Reunião e integração de equipe sobre o 2o trimestre	TODOS	SEDE	0%	seg, 2/11/2020	ter, 3/11/2020
5.1 - Expansão bibliográfica e modelagem dos modelos de redes neurais que realizam aplicação de séries temporais com dados exógenos, com base nas referências já catalogadas	PE	SEDE	0%	qua, 4/11/2020	qua, 18/11/2020
5.3 - Elaboração de metodologia de treino, reajuste empírico e avaliação dos modelos específicos de RNA para os negócios pesquisados	PE/PBI	SEDE	0%	qui, 19/11/2020	qui, 3/12/2020
5.4 - Elaboração de relatório científico dos modelos científicos	PE/TT3	SEDE	0%	sex, 4/12/2020	sex, 25/12/2020
7.1 - Estudo e modelagem da topologia do modelo global de rede neural artificial com base na topologia dos modelos específicos.	PE/TT3	SEDE	0%	seg, 28/12/2020	qui, 7/1/2021
9.1 - Elaboração e entrega de relatório científico final, com o estudo de todos os modelos de redes neurais	PE	SEDE	0%	sex, 8/1/2021	dom, 24/1/2021
9.2 - Entrega do relatório científico de todas as atividades do trimestre	PE	SEDE	0%	qua, 20/1/2021	dom, 31/1/2021
0.8 - Reunião e integração de equipe sobre o 3o trimestre	TODOS	SEDE	0%	seg, 1/2/2021	ter, 2/2/2021
11.1 - Expansão bibliográfica, estudo e modelagem de rede recorrente global aos negócios varejistas	PE/TT3	SEDE	0%	qua, 3/2/2021	sex, 5/3/2021
11.3 - Elaboração da metodologia de treino e avaliação da rede recorrente	PE/TT3	SEDE	0%	seg, 8/3/2021	qua, 7/4/2021
12.4 - Elaboração do relatório científico parcial, de todos os experimentos	PE/TT3	SEDE	0%	qui, 8/4/2021	qua, 21/4/2021
13.1 - Elaboração do relatório científico final da fase 1	PE/TT3	SEDE	0%	sex, 23/4/2021	qui, 29/4/2021
0.9 - Reunião de equipe e encerramento das atividades da fase 1	TODOS	SEDE	0%	sex, 30/4/2021	sex, 30/4/2021

6) Justifique a escolha do Nível para a bolsa (TT 1,2,3,4,4A,5 / PC 2,3,4).

Para as atividades do bolsista foi escolhido o nível TT4, graduado, especialista em Tecnologia de Informação (TI), com tema inerente à ciência de dados ou aprendizado de máquina para realização de atividades de apoio ao projeto de pesquisa, com a necessidade de experiência ou especialização técnica em aprendizado de máquina para poder realizar pesquisas e desenvolvimento dos modelos de aprendizado, bem como o estudo e expansão bibliográfica de modelos e suas topologias.

7) Justifique o Plano de Atividades em termos dos objetivos do Programa de Bolsas de Capacitação Técnica (apoio técnico ao projeto de pesquisa e treinamento técnico do bolsista).

As atividades do cronograma do bolsista TT4 correspondem aos objetivos do programa de capacitação técnica, dando treinamento técnico e capacitação técnica de pesquisa, desenvolvimento e aplicação da Inteligência Artificial, explorando toda a sua capacidade teórica, científica, e conectando a mesma com as necessidades do mercado que é muito promissor nesta área, e carece de estudos científicos e profissionais capacitados.

É importante ressaltar que o desafio científico do projeto não está no desenvolvimento de atividades de codificação, mas sim na pesquisa científica e na modelagem teórica dos modelos de aprendizado de máquina e que exigem conhecimentos acadêmicos especializados.

8) Informações adicionais

8.1) Observa-se que as bolsas TT4, TT4A e TT5 destinam-se a profissionais de TI (Tecnologia da Informação), para desenvolver atividades de TI (p.ex., desenvolvimento de software, integração de sistemas de software, projeto e análise de algoritmos, modelagem e implementação de bancos de dados complexos, arquitetura de software, projeto e implementação de sistemas embarcados e microcontrolados) junto ao projeto de pesquisa. As atividades de TI devem estar precisamente descritas. O tempo de experiência do candidato é contado após a conclusão da graduação.

8.2) Caso sejam solicitadas várias bolsas TT, os planos devem ser individuais, sem sobreposição de tarefas. Caso haja várias bolsas, é recomendada também a inclusão de um documento de apresentação descrevendo sucintamente como as bolsas se inter-relacionam e se complementam de forma a ficar claro a sua necessidade para a execução da pesquisa