

IA353A – Redes Neurais (FEEC/Unicamp)

Prof. Fernando J. Von Zuben

Notas do EF2

Observação: Como esta foi a última atividade avaliada, foram atribuídas notas até que o(a) aluno(a) atingisse um conceito final definitivo, selecionando uma sequência arbitrária de questões em cada relatório. Esta é a razão de algumas questões estarem sem nota, pois essas notas não influenciariam no resultado final de avaliação. Esta iniciativa permitiu que o professor realizasse um processo de avaliação mais abreviado, com a liberação das notas no prazo devido.

RA	Q5	Q6	Q7	Q8	Q9	Q10	Comentários
001658	1.0	1.0			1.0	1.0	
025809	1.0		1.0	0.9	1.0	0.5	Q8.3 (-0.1) Na formulação de EQM, é necessário deixar explícito o vetor de parâmetros livres da rede neural, ajustado durante o treinamento por reforço. Q10 (-0.5) A explicação da t-SNE está um tanto confusa. O fundamental é informar que se busca reproduzir no espaço de menor dimensão as distâncias relativas observadas no espaço original, usando distribuições diferentes. Não foi fornecido o notebook desta Q10. O PDF apresenta uma explicação muito concisa do que foi trabalhado na questão.
146244	0.7		1.0	0.8	1.0	0.9	Q5.1 e Q5.2 (-0.3) Faltaram as imagens reconstruídas e um passeio pelo manifold. Q10 (-0.1) Faltou falar das duas versões do Word2Vec. Q8.3 (-0.2) Não houve de fato uma explicação de como é definido o EQM.
156533	1.0			1.0	0.7	1.0	Q9.2 (-0.3) É válido o experimento, mas foi pedido que fosse para alguma outra base de imagens da literatura.
164381	1.0	1.0	1.0	0.85	1.0		Q8.3 (-0.15) A explicação para o EQM está sucinta demais.
166812	1.0	0.85	1.0	1.0	1.0	1.0	Q6.1.2 (-0.05) Faltou levantar alguma hipótese acerca do motivo pelo qual o preditor linear se baseia no valor atual da série para prever o próximo. Isso ocorre porque não há correlação com o passado da série e, assim, o próximo valor da série, quando comparado com o atual, tem aproximadamente o mesmo valor médio, excursionando para cima e para baixo, de forma descorrelacionada. Q6.1.3 (-0.05) O preditor linear, ao se basear no último valor da série, evita capturar tendências presentes apenas nos dados de

							treinamento (que são capturadas pelos modelos não-lineares, por serem mais flexíveis), o que permitirá que ele generalize melhor. Q6.1.4 (-0.05) Não é verdadeiro que o último valor da série temporal tem maior influência sobre os valores futuros, pois eles são descorrelacionados. O foco no offset está correto.
168116	0.9	0.86	1.0			1.0	Q5.1 (-0.1) Deveria ter mantido e ajustado o processo de visualização do manifold, com base nas trajetórias e no grid. Q6.2 (-0.14) Não implementou o item 4, mas apresentou propostas válidas. Houve desconto por atraso na entrega da Q6.
176625	1.0	0.95	1.0	1.0	1.0	1.0	Q6.2.3 (-0.05) Não implementou o preditor de múltiplos passos á frente, mas descreveu apropriadamente o seu fluxo de informação.
182716	1.0		1.0	0.85	1.0	1.0	Q8.3 (-0.15) A explicação para o EQM está sucinta demais.
203416	0.85		0.86		1.0	1.0	Q5.1 (-0.15) Mesmo sendo apenas uma comparação visual, a sua proposta não melhorou o resultado do notebook fornecido. Q7b (-0.14) Não indicou como foram definidos os hiperparâmetros da técnica.
211516	1.0		1.0	0.9	1.0	1.0	Q8.4 (-0.1) A técnica de experience replay não só armazena experiências vividas pelo agente, mas as utiliza para acelerar o aprendizado.
217002	1.0	0.86	0.8	0.7	1.0	0.6	Q5.2 Faltou trocar o caso de estudo, mas o uso do VAE é bem-vindo. Q6.1.2 (-0.06) Você não explicou por que a predição corresponde a uma versão atrasada da própria série. Q6.2.3 (-0.08) Não foi realizado. Q7.1 (-0.1) Faltou uma técnica. Q7.2 (-0.1) Fez uma busca, mas não disse qual a parametrização adotada no último resultado. Q8 (-0.3) Apresentou os Q-valores apenas para dois casos, sem identifica-los (foram solicitados 3 casos representativos, devidamente identificados). Para explicar como é definido o EQM, é necessário recorrer à equação de Bellman. Falar que EQM envolve um erro elevado ao quadrado é redundante com a própria sigla. Não se justificam 2 notebooks para cada labirinto. Não apresentou a trajetória e os Q-valores do caso 2b. Q10 (-0.4) Faltou falar das duas versões do Word2Vec. A descrição do t-SNE está muito pobre. As cores usadas não têm a ver com o t-SNE e sim com a aplicação dele nesta atividade. A interpretação dos resultados obtidos em todo o notebook está incompleta.
228101							Não foi entregue.
229060	1.0	1.0			1.0	1.0	

229960	1.0		1.0	1.0	1.0	1.0	
232881	1.0	0.82	1.0		1.0		<p>Q6.1.2 (-0.05) Faltou levantar alguma hipótese acerca do motivo pelo qual o preditor linear se baseia no valor atual da série para prever o próximo. Isso ocorre porque não há correlação com o passado da série e, assim, o próximo valor da série, quando comparado com o atual, tem aproximadamente o mesmo valor médio, excursionando para cima e para baixo, de forma descorrelacionada. Q6.1.3 (-0.05) O preditor linear, ao se basear no último valor da série, evita capturar tendências presentes apenas nos dados de treinamento (que são capturadas pelos modelos não-lineares, por serem mais flexíveis), o que permitirá que ele generalize melhor. Q6.2.4 (-0.08) Você pode diferenciar a série até obter a estacionariedade ou então adotar um preditor que aprende o tempo todo, enquanto faz as previsões.</p>
234613	0.85	0.73	0.6	0.6	1.0	0.6	<p>Q5 (-0.15) Só tem o notebook referente à Parte 1 e não o da Parte 2. O caminho ao longo do manifold não deveria incluir tanta porção onde não há dados. Alguns resultados da Parte 2 estão no PDF, indicando que se usou um VAE. Deveria ter alterado o caso de estudo. Q6.1.2 (-0.08) A sua resposta não está correta. Uma análise direta dos coeficientes do preditor linear permite constatar que ele está focado basicamente no último valor da série temporal. Q6.2.2 (-0.04) Está correto falar da flexibilidade do modelo, mas não tem a ver com a frequência de oscilação da série temporal. A não-linearidade da série não é representável pelo modelo linear. Q6.2.3 (-0.08) Não realizado. Q6.2.4 (-0.04) Não está claro como utilizar o conceito de média móvel. Q6.3 (-0.03) Por que o preditor linear tem tantos coeficientes assim? O preditor linear está acompanhando a série, mas com um problema de offset que precisava ser melhor esclarecido. Q7a Não haverá desconto, mas os resultados do arquivo PDF deveriam estar no notebook também. Q7b (-0.4) As imagens apresentadas não correspondem aos 10 dígitos, mas a uma mesma imagem com parametrizações distintas do algoritmo. Novamente, os resultados do PDF deveriam estar no notebook. Q8.3 e Q8.4 (-0.4) Não foram feitas. Q9 Faltou o notebook com os resultados para a base MNIST, os quais foram apresentados apenas no PDF. Q10 (-0.4) Está incompleta a definição de t-SNE, não evidenciando a técnica e o que ela busca realizar. Também faltaram muitos conceitos relevantes para o Word2Vec.</p>
234617	1.0	1.0	1.0	1.0	1.0	1.0	
234619	0.95		1.0		1.0	1.0	<p>Q5 (-0.05) As trajetórias pelo manifold estão muito locais e uma delas não levou ao</p>

							resultado.
234633	1.0	1.0	1.0	0.9	1.0	0.65	Q8 (-0.1) Para explicar como é definido o EQM, é necessário recorrer à equação de Bellman. Q10 (-0.35) Está incompleta / não plenamente convincente a definição de t-SNE, não evidenciando a busca por gradiente e o conceito de perplexidade. O kernel t-student não mapeia as amostras, mas é utilizado para tentar reproduzir, no espaço de dimensão reduzida, a posição relativa dos dados no espaço original. Também faltaram muitos conceitos relevantes para o Word2Vec: o fato de os mapeamentos envolvidos serem lineares, a codificação se dar no gargalo do autoencoder-like, o qual representa um espaço denso de dimensão apropriada (entre 50 e 500). Podia ter dado o nome das duas variações: CBOW e Skip-Gram.
235624	1.0		1.0	1.0	1.0		
258499			1.0		1.0	1.0	
261152	1.0			1.0	1.0	1.0	
261663	1.0		1.0	1.0	1.0	0.85	Q10 (-0.15) Está incompleta a definição da técnica t-SNE, pois faltou apresentar como ela faz o que você afirma que é feito por ela.
262249	0.8	1.0	1.0	1.0	1.0	0.75	Q5.1 (-0.2) Se esperava um ganho de desempenho sem precisar aumentar a dimensão do manifold. Q10 (-0.25) Que tarefa de PLN o Word2Vec está resolvendo? A definição de t-SNE está incompleta. É fundamental falar das duas funções de probabilidade envolvidas.
265568	1.0		1.0	1.0	1.0	1.0	
265945	1.0	0.9	1.0	1.0			Q6.2.4 (-0.1) Está correto propor transformar a série não-estacionária em uma série estacionária, mas não da forma como você sugere. Você pode diferenciar a série até obter a estacionariedade ou então adotar um preditor que aprende o tempo todo, enquanto faz as previsões. Q6.3 É esperado que o preditor não-linear apresente um desempenho um pouco superior àquele do preditor linear para a série escolhida, de acordo com comparações já realizadas por mim. Eu não consegui detectar problemas na sua metodologia, no entanto.