

IA353 – Redes Neurais (1s2020)

Roteiro de Estudos (Parte 2)

Índice Geral

1	Notas de Aula – Tópico 5	2
2	Notas de Aula – Tópico 6	8
3	Notas de Aula – Tópico 7	12

1 Notas de Aula – Tópico 5

1. Qual é o princípio da computação analógica?
2. Como se dá o processo de simulação de um neurocomputador via computador digital?
3. Há alguma chance de circuitos analógicos sustentarem a IA no futuro? Nota: Uma discussão interessante pode ser encontrada em: <https://semiengineering.com/using-analog-for-ai/>
4. Por que se pode afirmar que o computador digital é a mais determinística das máquinas concebidas pela mente humana?
5. O que é estado de um sistema dinâmico? O que é espaço de estados? O que é uma trajetória no espaço de estados?
6. Procure definir dinâmica usando o conceito de estado de um sistema dinâmico.

7. Toda dinâmica é determinística, estacionária, linear e contínua? Ou existem 16 tipos de dinâmicas diferentes, considerando a presença ou não desses 4 atributos? Faça essa análise considerando que o estado do sistema dinâmico é contínuo.
8. Usando esses 4 atributos da dinâmica de um sistema, procure definir a dinâmica de um computador digital.
9. Com base na resposta da questão anterior, explique como um computador digital pode simular um sistema com dinâmica contínua e estado contínuo.
10. Sendo uma máquina determinística (em condições normais de operação), é possível gerar números aleatórios em computador digital? Nota: Favor consultar <https://www.iro.umontreal.ca/~lecuyer/myftp/papers/handstat.pdf>
11. Apresente as equações de estado e de saída para um sistema dinâmico não-linear em sua representação por espaço de estados, tanto para a dinâmica contínua como para a dinâmica discreta.

12. Dadas as representações da questão anterior, explique o que é linearizar um sistema dinâmico em torno de um ponto de operação.
13. O que é um sistema dinâmico autônomo?
14. Apresente ao menos dois comportamentos que podem ser exibidos por sistemas dinâmicos autônomos não-lineares e não podem ser exibidos por sistemas dinâmicos autônomos lineares.
15. Quais são os 4 comportamentos em regime que podem ser exibidos por sistemas dinâmicos (seja para dinâmica discreta ou contínua) autônomos não-lineares?
16. Um mesmo sistema dinâmico autônomo não-linear pode exibir todos esses 4 comportamentos em regime, dependendo do estado inicial?
17. Um certo sistema dinâmico autônomo não-linear pode exibir apenas um desses 4 comportamentos em regime, independente do estado inicial?
18. O que é bacia de atração?
19. O que é um atrator estranho?

20. Procure definir sensibilidade à condição inicial (efeito borboleta) em sistemas caóticos.
21. O que é um retrato de fase?
22. O que é bifurcação do comportamento dinâmico de sistemas autônomos não-lineares?
23. O que é um diagrama de teia (*cobweb*)?
24. Por que se diz que a rede de Elman é um modelo universal para sistemas dinâmicos lineares?
25. O que é uma rede neural recorrente? Há formas distintas de se produzir uma rede neural recorrente ou elas sempre exibem o mesmo tipo de arquitetura?
26. Apresente ao menos dois fatores para sustentar que treinar uma rede neural recorrente é mais desafiador do que treinar uma rede neural não-recorrente.
27. Explique como funciona a técnica de *backpropagation through time*.

28. Por que gradientes que explodem e que se anulam são indesejados durante o treinamento de uma rede neural (seja ela recorrente ou não-recorrente)?
29. Que propriedade de um sistema dinâmico autônomo não-linear foi explorada por Hopfield para implementar memórias endereçáveis por conteúdo?
30. Defina a rede neural de Hopfield.
31. Quais são os 4 principais fatores que motivam a implementação de redes neurais em hardware?
32. Explique como a dinâmica de relaxação de um sistema dinâmico autônomo não-linear pode ser utilizada para implementar memória endereçável por conteúdo.
33. O que se pode dizer de um sistema dinâmico autônomo não-linear para o qual é possível definir uma função de Lyapunov.
34. Dadas as memórias fundamentais a serem armazenadas por uma rede neural de Hopfield, como é possível definir os pesos sinápticos?

35. Como é possível encontrar algebricamente pontos de equilíbrio em sistemas dinâmicos autônomos? Considerar dinâmica contínua e discreta.
36. O que são atratores espúrios?
37. Comente sobre a capacidade de memória de uma rede de Hopfield.
38. Comente sobre a dificuldade de se armazenar memórias muito parecidas (que estão longe de serem ortogonais entre si).
39. Explique como Hopfield adaptou a sua rede neural recorrente para resolver o problema do caixeiro viajante.
40. Defina redes neurais com estados de eco.
41. Por que as redes neurais com estados de eco são consideradas uma vertente de *reservoir computing*?
42. Defina a propriedade de estados de eco e procure apresentar uma metodologia para promover, na prática, a sua ocorrência.

43. Em que as redes neurais com estados de eco se assemelham às máquinas de aprendizado extremo (ELM, do inglês *Extreme Learning Machines*)?
44. Por que, no caso de uma rede neural recorrente com entrada externa e dinâmica estacionária, dada a mesma entrada externa em dois instantes distintos de tempo, a saída pode ser diferente?

2 Notas de Aula – Tópico 6

45. Faça uma distinção entre os conceitos de ordem e organização.
46. É possível caracterizar completamente um sistema complexo com base no pleno conhecimento de seus elementos constituintes?
47. Liste as três condições necessárias para a ocorrência de auto-organização.
48. Apresente exemplos práticos da ocorrência dessas 3 condições em sistemas auto-organizados do mundo real.
49. Dê exemplos de fenômenos espaço-temporais do mundo real.

50. Considerando objetivos especificados a priori (por exemplo, de como um sistema deve se comportar no futuro), quais são os principais desafios quando se busca influenciar (por atuação externa) a evolução de um sistema complexo visando maximizar esses objetivos?
51. Por que a definição de múltiplos protótipos para uma certa distribuição de dados não pode ser resolvida de forma fechada, por técnicas de otimização, requerendo a adoção de mecanismos de auto-organização?
52. Explique como opera o algoritmo *k-means*.
53. Apresente uma explicação didática que demonstre que o algoritmo *k-means*, em sua forma original, está sujeito a mínimos locais. Em seguida, proponha uma heurística que reduza a chance de se obter mínimos locais ruins.
54. Se você fosse propor um índice de qualidade de agrupamento, não importando sua originalidade, que informações você considera que esse índice deve contemplar?

55. Quais são e por que são definidas duas distâncias entre os neurônios de um mapa de Kohonen?
56. Por que se diz que um mapa de Kohonen, embora também esteja sujeito a mínimos locais, tende a preservar maximamente as relações de distância existente no espaço original dos dados usados no treinamento?
57. Quais são as vantagens e as desvantagens quando se comparam mapas de Kohonen 1D e 2D?
58. Qual a diferença entre um mapa de Kohonen plano e um toroidal?
59. Por que geralmente o tamanho da vizinhança e o passo de ajuste são reduzidos ao longo do processo de auto-organização?
60. Apresente um pseudo-código que descreva o processo de treinamento não-supervisionado do mapa auto-organizável de Kohonen, o qual, após o treinamento, tende a ativar mais intensamente neurônios próximos no mapa sempre que a entrada recebe estímulos parecidos.

61. Por que, quando computacionalmente viável, é recomendado que a inicialização dos pesos do mapa de Kohonen 2D siga o hiperplano formado pelos dois componentes principais dos dados disponíveis para treinamento?
62. Descreva como obter a matriz-U e como a utilizar para a discriminação de agrupamentos.
63. Além da discriminação de agrupamentos, que outras aplicações relevantes você apontaria para mapas de Kohonen?
64. Como um mapa de Kohonen 1D com plasticidade (que pode ganhar ou perder neurônios e arestas) pode ser empregado na busca por uma boa solução para o problema do caixeiro viajante?
65. Por que não é elementar inserir plasticidade num mapa de Kohonen 2D, seja com vizinhança retangular ou hexagonal?

66. Procure se familiarizar com o mapa auto-organizável construtivo *Growing Neural Gas*, explicando ao menos como funciona e para que serve a idade de cada aresta.
67. O que é *Learning Vector Quantization*?

3 Notas de Aula – Tópico 7

68. Se uma máquina de aprendizado obteve um desempenho aceitável durante o treinamento, o que pode justificar uma queda inaceitável de desempenho quando em operação de campo (ou junto a dados de teste)?
69. Por que poucos dados, dados ruidosos e dados de elevada dimensão, conjuntamente ou de forma isolada, representam desafios para máquinas de aprendizado?
70. Procure na literatura pelo conceito de maldição da dimensionalidade (*curse of dimensionality*) e o vincule ao aprendizado a partir de dados.
71. Explique o dilema bias-variância.

72. Visto que em certas aplicações mais complexas foi demonstrado que dados importam mais que modelos de aprendizado, consulte a literatura e apresente ao menos 3 técnicas voltadas para *data augmentation*, por exemplo, consultando: <https://towardsdatascience.com/data-augmentation-for-deep-learning-4fe21d1a4eb9>
73. Para a regularização de mapeamentos unidimensionais, além das sugestões presentes nas notas de aula, existem LOESS (*locally estimated scatterplot smoothing*) e LOWESS (*locally weighted scatterplot smoothing*). Procure entender como funcionam essas técnicas de regressão não-paramétrica.
74. Explique como funciona o truque do *kernel*, fundamento da regressão por função *kernel* e das máquinas de vetores-suporte.
75. Explique a diferença entre *ridge regresson* e LASSO.
76. Explique como LASSO pode ser usado para seleção de variáveis.
77. O que é *elastic net*?
78. Qual a diferença entre risco esperado e risco empírico?
79. Defina a dimensão Vapnik-Chervonenkis.

80. No caso de um espaço de hipóteses de cardinalidade infinita, o que é necessário para que o risco empírico se aproxime do risco esperado?
81. Explique o princípio da minimização do risco estrutural.
82. Defina hiperplano ótimo num problema de classificação binária.
83. Por que o hiperplano ótimo tende a maximizar a capacidade de generalização, quando comparado com qualquer outro hiperplano?
84. O que são vetores-suporte?
85. Como se define o número de vetores-suporte?
86. Por que o espaço de características geralmente tem dimensão maior que o espaço de entrada, particularmente quando o problema de classificação é não-linear?
87. Como lidar com problemas não-linearmente separáveis ou com erros de classificação, seja no espaço original, seja no espaço de características?
88. Como escolher a função *kernel* dadas diversas funções candidatas? E se a função *kernel* tiver parâmetros a determinar, como ajustá-los?

89. Como lidar com problemas de classificação que envolvem mais de duas classes?
90. Como estender a formulação de SVM para resolver problemas de regressão?
91. Em SVM, a dimensão VC está associada a quê?
92. Descreva o princípio de inferência transdutiva.
93. Na extensão de SVM para clusterização, o que é equivalente ao hiperplano ótimo?