Intelligent systems:
sample questions for oral exams in 2020/21

Nature-inspired computing

1. Describe the main component of an evolutionary program: population representation, generation, selection, combination, replacement, and stopping criteria?
2. Describe when to use genetic algorithms?
3. Describe the strengths and weaknesses of evolutionary programs.
4. Describe the main characteristics of genetic algorithms (GA) and genetic programming (GP).
5. Describe terms from evolutionary computation such as population variability, fitness function, coevolution.
6. Describe different gene representations in GA, operations on them, and their strengths and weaknesses: bit and numeric vectors, strings, permutations, trees.
7. What are linear crossover, Gray coding of binary numbers, adaptive crossover, gaussian mutation, Lamarckian mutation, and elitism? What are their advantages compared to baselines?
8. Describe the following evolutionary models: proportional and rank proportional roulette wheel, tournaments, single tournament, and stochastic universal sampling?
9. How to prevent niche specialization in GA?
10. Explain hypotheses on why GAs work?
11. What are the typical parameters of GAs?
12. Where to use GAs and where not?
13. Why are GAs suitable for multiobjective optimization, and what is Pareto optimal solution?
14. Explain the main problems of genetic programming.

Machine learning (ML)

15. Describe the two main goals of ML, prediction and inference, and explain why they are sometimes in contradiction.
16. What parametric and non-parametric ML methods exist?
17. Describe the main characteristics of supervised, unsupervised, and semi-supervised ML methods?
18. What is the difference between regression and classification? Give examples of problems for each type.
19. What are association rules, and how they differ from decision rules?
20. What are outliers in ML?
21. Contrast two different views on ML: as optimization and as search.
22. Describe different properties of ML models: bias, variance, generalization, hypothesis language.
23. What is the bias-variance trade-off in ML?
24. Describe the double descent concerning bias-variance trade-off.
25. Describe bias-variance trade-off in relation to kNN classifier.

26. Describe methods that can speed-up the kNN algorithm: k-d trees, R-trees, RKD-tree, locally sensitive hashing, and hierarchical k-means.
27. What are the Bayes error rate and Bayes optimal classifier?
28. Describe properties of the following models: kNN, decision rules, bagging, boosting, random forests, stacking, MARS, AODE, SVM, neural networks.
29. What is the difference between training and testing error? Why we need an evaluation set?
30. Describe the properties and purpose of evaluation with cross-validation. Describe different biases of ML models stemming from data: reporting bias, automation bias, selection bias, group attribution bias, implicit bias.
31. What is the no-free-lunch theorem?
32. Describe three types of feature selection methods: filter, wrapper, and embedded methods. What are the main differences between them?
33. Describe the difference between impurity based and context-sensitive attribute evaluation.
34. Describe the main ideas of information gain and ReliefF evaluation measure.
35. Explain how regularization can be used as a feature selection method?
36. Describe ridge regression and lasso and the difference between them?
37. What are the advantages and disadvantages of the wrapper method for feature selection?
38. Describe the confusion matrix and evaluation measures based on it?
39. Describe ROC curves, sensitivity, specificity, precision, recall, F-measure, classification accuracy, mean squared error.
40. What are the ideas of unsupervised and semi-supervised feature selection?
41. How can we increase the stability of feature selection?
42. Describe the main idea s of multi-view, multi-label, and multitask learning.
43. What do online learning and online feature selection mean?
44. Explain the main ideas of ensemble methods in ML, why and when they work?
45. Explain the main differences between bagging and random forests?
46. What is the out-of-bag error estimation?
47. How can one evaluate attributes with random forests or produce a similarity matrix?
48. Describe the main parameters of random forests and boosting?
49. Describe the main idea of gradient boosting?
50. Describe the notion of margin in kernel methods.
51. What is the purpose of different kernels (linear, polynomial, RBF) in SVM?
52. Describe how to use SVM for more than two classes?
53. Describe different activation functions in neural networks (NNs).
54. Describe the main idea of backpropagation learning for NNs.
55. Describe the role of criterion (loss) function in NN?
56. Describe the strengths and weaknesses of NN?
57. Describe a few techniques for overfitting prevention in NNs.
58. What are deep neural networks? What are their main strengths and weaknesses?
59. What are the recurrent networks?
60. Describe the convolutional neural networks (CNN).
61. Describe different components of CNNs.

62. What are the advantages and disadvantages of CNNs?
63. What is 1d and 2d convolution?
64. Describe the main idea and components of autoencoders?
65. What is denoising autoencoder?
66. Describe the main idea and components of the generative adversarial networks?
67. Describe different inference methods for predictive methods.
68. Describe different techniques for the explanation of predictions.
69. What is the role of clustering in interpretability?
70. Describe the main idea of perturbation-based explanation methods?
71. Explain the difference between instance-based and model-based explanations?
72. Explain the main idea of the IME, LIME, and SHAP explanation technique?

Natural language processing (NLP)

73. What is the Turing test?
74. What is the micro-world approach to NLP?
75. Describe the stages of linguistic analysis?
76. Describe how to preprocess text in text mining.
77. Describe lemmatization, stemming, POS tagging, dependency parsing, and named entity recognition.
78. Describe the basic language resources for English and Slovene (or your language).
79. Describe the structure of WordNets.
80. Describe approaches to document retrieval.
81. Describe the inverted file index.
82. Compare search with logical operators and ranking based search.
83. Describe one-hot-encoding and bag-of-words representation.
84. Describe how to use term-document and term-term matrix?
85. What is word embedding? Which embeddings are sparse and which are dense?
86. Describe the use of cosine similarity on documents.
87. Describe TF-IDF weighting.
88. Describe precision, recall, and F_1 measures in document retrieval.
89. Describe problems of web search and possible improvements.
90. Describe the idea of the PageRank algorithm and its possible uses.
91. Describe the main ideas and implementation of LSA, word2vec, ELMo, and BERT.
92. Which are the desired properties of word embeddings?
93. Compare different types of word embeddings.
94. Describe a few relations expressed with modern word embeddings.
95. What sort of biases are reflected in word embeddings?
96. How to use BERT and multilingual BERT for text classification?
97. Describe the idea and a few uses of cross-lingual embeddings?
98. Describe a few semantic technologies and a few important NLP tasks.

99. How to approach text summarization, sentiment classification, machine translation (MT), or question answering problems?
100. What are the language model and translation model in MT?
101. What is the encoder-decoder model in NLP?
102. What is the attention mechanism in deep neural networks?

Reinforcement learning (RL)

103. Describe when and why to apply RL.
104. What are the differences between supervised learning and RL?
105. Describe the explore or exploit dilemma in RL?
106. Describe the four main components of RL and their role.
107. How the interface between the agent and environment works in RL?
108. Describe returns for episodic and continuing tasks.
109. What is the discounted return, and what is its role?
110. What is the average reward model, and what are its advantages and disadvantages?
111. What is the role of Markov property in RL?
112. Describe the Markov decision problem (MDP).
113. What sort of learning simplifications does MDP allow in RL?
114. Describe the value function and action-value functions?
115. Describe the Bellman equations and their role in RL?
116. What is the role of the optimal value function and optimal action-value function?
117. How can we get the optimal policy from the optimal action-value function?
118. How to solve Bellman optimality equations?
119. When and how dynamic programming is used in RF?
120. Describe policy-value iteration, value iteration, and policy iteration approaches to RF?
121. Describe the convergence criterion for value iteration.
122. Describe the Monte Carlo approach to RL and when it is used.
123. Describe the ε-greedy policy.
124. Describe learning with time differences (TD) in RL?
125. Describe the Q-learning.
126. What are the updates in Q-learning? How to assure exploration?
127. How to use function approximation in RL?
128. How to measure and compare the learning performance of RL learners?