ourdiu.com | Shared by SUPAN ROY

Daffodil International University

Faculty of Science & Information Technology

Department of Software Engineering

Final Examination, Summer 2026

Course Code: DS 423; Course Title: Machine Learning Driven Data Analysis II and Communicating Data Insights

Sections & Teachers: A,B,C (JJS)

Time: 2:00 Hrs **Marks: 40**

Answer ALL Questions

[The figures in the right margin indicate the full marks and corresponding course outcomes. All portions of each question must be answered sequentially.]

1.	*a)*	Professor Severus Snape is developing an AI system to automatically identify different magical potions from their images. To ensure high classification accuracy, he designs a deep convolutional neural network with the following characteristics: • Every convolutional layer uses only 3×3 filters instead of larger filters such as 5×5 or 7×7. • A 2×2 max-pooling layer is placed after each convolutional block to reduce the image size while preserving important information. • After extracting all features, the network uses three fully connected layers to determine the correct potion type. **Evaluate** the suitability of the CNN architecture for Professor Snape's potion classification system by identifying the architecture, briefly explaining three key theoretical contributions, and discussing four limitations of this architecture.	**[Marks-5]**	**CLO-2** *Level-V*
	b)	The Ministry of Magic is developing an AI system to recognize magical spells from images captured during wizard duels. Some spells produce very small visual effects (such as tiny sparks), while others create large magical patterns. Instead of using only one filter size, the researchers design a CNN that processes the same input feature map through multiple branches simultaneously and the outputs from all branches are concatenated into a single feature map before being passed to the next layer. **Assess** the suitability of the CNN architecture for this spell recognition system by identifying the architecture, drawing the architecture of a single module, briefly explaining how the module recognizes both small and large magical spell patterns, and writing two key theoretical contributions and three advantages of this architecture.	**[Marks-5]**	
2.	*a)*	At Imperial College of Engineering, Rancho develops an AI system to predict whether a student will successfully complete a complex engineering project. The system analyzes a sequence of information collected over several weeks, including quiz scores, assignment marks, attendance, project progress, and laboratory performance. Initially, Rancho uses a standard RNN. However, he notices that when the sequence becomes long, the model gradually forgets important information from the earlier weeks, resulting in poor predictions. To solve this problem, he replaces the model with another recurrent architecture that can decide what information should be remembered, what should be discarded, and what new information should be stored using multiple gating mechanisms and an internal memory.	**[Marks-6]**	

ourdiu.com | Shared by SUPAN ROY

	Evaluate the suitability of the final architecture for Rancho's student performance prediction system by identifying the architecture, explaining its architecture and working mechanism with the help of its gating mechanisms, and describing how this architecture reduces the vanishing-gradient problem.		**CLO-3** *Level-V*
b)	Inspired by Rancho's work, Chatur develops another AI system to analyze sequences of students' computer usage logs to detect malware infections in the college laboratory. Since the system must run efficiently on computers with limited computational resources, Chatur selects a recurrent architecture that combines the memory and hidden state into a single state and uses only two gates instead of three. This design reduces the number of parameters while still preserving long-term information. **Compare** the suitability of the recurrent neural network architecture for Chatur's malware detection system by identifying the architecture and writing five differences between the architecture identified in Question 2(a) and the architecture identified in Question 2(b)	**[Marks-4]**	
c)	Assume the LSTM cell state has dimension 2. The previous hidden state, previous cell state, and current input are: $$h_{t-1} = \begin{bmatrix} 0.2 \\ 0.5 \end{bmatrix}, \quad C_{t-1} = \begin{bmatrix} 0.4 \\ -0.3 \end{bmatrix}, \quad x_t = \begin{bmatrix} 0.7 \\ 0.9 \end{bmatrix}$$ The weight and bias matrix for forget, input and output gates are respectively: $$W_f = \begin{bmatrix} 0.3 & -0.4 & 0.5 & 0.2 \\ -0.2 & 0.6 & 0.1 & -0.5 \end{bmatrix}, \quad b_f = \begin{bmatrix} 0.10 \\ -0.05 \end{bmatrix} \quad W_i = \begin{bmatrix} 0.4 & -0.1 & 0.2 & 0.3 \\ 0.1 & 0.5 & -0.3 & 0.4 \end{bmatrix}, \quad b_i = \begin{bmatrix} 0.00 \\ 0.05 \end{bmatrix}$$ $$W_C = \begin{bmatrix} 0.2 & 0.3 & 0.4 & -0.2 \\ -0.5 & 0.1 & 0.6 & 0.2 \end{bmatrix}, \quad b_C = \begin{bmatrix} 0.05 \\ 0.00 \end{bmatrix} \quad W_o = \begin{bmatrix} 0.5 & -0.2 & 0.3 & 0.1 \\ -0.3 & 0.4 & 0.2 & 0.5 \end{bmatrix}, \quad b_o = \begin{bmatrix} -0.05 \\ 0.10 \end{bmatrix}$$ **Assess** the LSTM cell computation by calculating the final hidden state. Show all intermediate calculations and report values rounded to three decimal places.	**[Marks-10]**	
3. *a)*	Welton Academy develops an artificial intelligence system called "Carpe Diem AI." The system analyzes classroom conversations between Mr. Keating and his students and generates thoughtful responses. For example, the system receives the sentence: "Mr. Keating encourages the students to view poetry from a different perspective." The AI must understand the relationships among words such as "Keating," "students," "poetry," and "perspective." Instead of processing words one after another, it processes the entire sentence simultaneously while using positional encoding to preserve word order. During text generation, it ensures that each newly generated word cannot access future words. **Analyze** the deep learning architecture used in the Carpe Diem AI system. Draw its complete architecture, clearly labeling all major components. Then explain how the Multi-Head Self-Attention mechanism works in this architecture.	**[Marks-5]**	**CLO-4** *Level-IV*
b)	During one attention head, the word "poetry" attends to three words in the sentence. The Query, Keys, and Values are given below. $$Q = [2 \quad 1]$$ $$K = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 2 \end{bmatrix} \qquad V = \begin{bmatrix} 4 & 1 \\ 2 & 3 \\ 5 & 2 \end{bmatrix} \qquad d_k = 2$$ Using the Scaled Dot-Product Attention equation, **Examine** the given values and calculate the attention output for the query vector.	**[Marks-5]**	