🔑 Must-Know Keywords & Simpler One-Line Answers

பரீட்சையில் விடை எழுதும்போது இந்த வார்த்தைகள் உங்கள் பதில்களில் கண்டிப்பாக இடம் பெற வேண்டும்.

Section A: AI Basics & Agent Architectures

- **Artificial Intelligence (AI):** மனிதனைப் போல சிந்தித்துச் செயல்படும் கணினி அமைப்புகளை உருவாக்குவது.
- **Turing Test:** ஒரு இயந்திரத்தின் புத்திசாலித்தனத்தை மனித அறிவாற்றலுடன் ஒப்பிட்டு அளவிடும் சோதனை.
- **Intelligent Agent:** தன் சூழலை உணர்ந்து (Sensors), இலக்கை அடையச் சுயாதீனமாகச் செயல்படும் (Actuators) ஓர் அமைப்பு.
- **Agent Function:** சூழலில் இருந்து பெறப்படும் தகவல்களை (Percepts), தகுந்த செயல்களாக (Actions) மாற்றித் தரும் கணித வரைபடம்.
- **Sensors:** கேமரா, ரேடார் போல சுற்றுச்சூழலின் மாற்றங்களை உணர்ந்து சிக்னல்களாக மாற்றும் கருவிகள்.
- **Actuators:** பிரேக், சிக்னல் விளக்குகள் போல சிக்னல்களைப் பெற்று உடல் ரீதியான இயக்கத்தை ஏற்படுத்தும் கருவிகள்.
- **PEAS:** ஒரு ஏஜென்ட்டின் தன்மையை வரையறுக்கும் முக்கிய அளவுகோல் (Performance measure, Environment, Actuators, Sensors).

Section B: Machine Learning Core Concepts

- **Supervised Learning:** லேபிளிடப்பட்ட (Labeled) பயிற்சித் தரவுகளைக் கொண்டு மாடல்களுக்குக் கற்றுக்கொடுக்கும் முறை.
- **Unsupervised Learning:** லேபிள்கள் இல்லாத தரவுகளில் உள்ள மறைமுகத் தொடர்புகளை (Patterns) மாடல் தானாகவே கண்டறியும் முறை.
- **Reinforcement Learning:** தவறுகளுக்குத் தண்டனையும் (Punishments), சரியான செயல்களுக்குப் புள்ளிகளும் (Rewards) வழங்கி சோதனைகள் மூலம் ரோபோக்களைக் கற்க வைக்கும் முறை.
- **Classification:** தரவுகளை வெவ்வேறு பிரிவுகளாக/வகைகளாக வகைப்படுத்துவது (எ.கா: Spam அல்லது Not Spam).
- **Regression:** தரவுகளைக் கொண்டு தொடர்ச்சியான எண்களைக் கணிப்பது (எ.கா: வீட்டின் விலை, வெப்பநிலை).

Section C: Data Preprocessing & Evaluation

- **Data Cleaning:** தரவுகளில் விடுபட்ட மதிப்புகள் (Missing values) மற்றும் டூப்ளிகேட்களை நீக்கிச் சுத்தப்படுத்துவது.
- **Normalization:** வெவ்வேறு அளவுகளில் (Scales) இருக்கும் எண்களை 0 முதல் 1 என்ற குறிப்பிட்ட எல்லைக்குள் மாற்றுவது.
- **Confusion Matrix:** ஒரு வகைப்பாட்டு மாடலின் (Classification Model) உண்மையான மற்றும் கணிக்கப்பட்ட முடிவுகளை ஒப்பீட்டுக் காட்டும் அட்டவணை.
- **Accuracy:** மாடல் கணித்த மொத்த முடிவுகளில், எத்தனை முடிவுகள் மிகச் சரியானவை என்பதன் ஒட்டுமொத்த சதவீதம்.
- **Precision:** மாடல் "பாசிட்டிவ்" என்று கணித்தவற்றில், உண்மையில் எத்தனை சதவீத முடிவுகள் சரியாக உள்ளன என்ற அளவு.

- **Recall:** நிஜ உலகத் தரவில் உள்ள மொத்த பாசிட்டிவ் மதிப்புகளில், மாடல் எத்தனை சதவீதத்தைச் சரியாகக் கண்டுபிடித்து மீட்டெடுத்துள்ளது என்ற அளவு.

Section D: Deep Learning & Generative Models

- **Neural Network:** மனித மூளையின் நியூரான்களைப் போல அடுக்குகளாகச் செயல்படும் ஒரு செயற்கை கணித அமைப்பு.
- **Activation Function:** நியுறல் நெட்வொர்க்கிற்கு நேரியல் அல்லாத தன்மையைக் (Non-linearity) கொடுத்து, சிக்கலான வடிவங்களைப் புரிந்துகொள்ள உதவும் சார்பு.
- **ReLU:** எதிர்மறை எண்களை (Negative values) 0 ஆக மாற்றும் ஒரு எளிய வேகமான Activation Function.
- **Sigmoid:** எந்தவொரு எண்ணையும் 0 மற்றும் 1 க்கு இடைப்பட்ட சாத்தியக்கூறு (Probability) மதிப்பாக மாற்றும் Activation Function.
- **Epoch:** கொடுக்கப்பட்ட பயிற்சித் தரவை மாடல் முழுமையாக ஒருமுறை படித்துப் பழகும் ஒரு சுற்று (One complete pass).
- **Loss Function:** மாடல் கணித்த பதிலுக்கும், உண்மையான பதிலுக்கும் இடையே உள்ள பிழையை (Error) அளவிடும் அளவுகோல்.
- **Backpropagation:** Loss Function காட்டும் பிழையை வைத்து, நியூரல் நெட்வொர்க்கின் கணித மதிப்புகளை (Weights) பின்னோக்கிச் சென்று மாற்றி அமைக்கும் கற்றல் வழிமுறை.
- **Overfitting:** மாடல் பாடங்களைப் புரிந்து கொள்ளாமல் தரவுகளை அப்படியே மனப்பாடம் செய்வதால், புதிய தரவுகளில் மோசமாகச் செயல்படும் ஒரு சிக்கல்.
- **Generative Model (GAN):** உலகிலேயே இல்லாத புதிய தரவுகளையும் (படங்கள்/வீடியோ) அச்சு அசலாகப் புதிதாக உருவாக்கும் AI தொழில்நுட்பம்.

📝 5 Essay Type Questions: Detailed Guidance

உங்களது ஸ்கிரீன்ஷாட்டில் உள்ள ஐந்து தலைப்புகளின் அடிப்படையிலேயே மாதிரி வினாத்தாளும் அமைந்துள்ளது. ஒவ்வொரு கேள்வியையும் பரீட்சையில் எவ்வாறு அணுக வேண்டும் என்பதற்கான முழு வழிகாட்டல் இதோ:

💡 Q1: Introduction to AI & PEAS (Smart Traffic System)

- அணுகுமுறை: இந்த வினா முழுமையாக தத்துவார்த்த ரீதியானது (Theoretical).
- வழிகாட்டல்:
  - **Intelligent Characteristics:** போக்குவரத்து மேலாண்மை அமைப்பு எவ்வாறு தானாகவே சிக்னல் நேரத்தை மாற்றுகிறது என்பதை எழுதி, அதற்குப் பின்னால் உள்ள **Perception** (கேமரா மூலம் உணர்தல்) மற்றும் **Autonomy** (சுயாதீன முடிவு எடுத்தல்) ஆகிய பண்புகளை விளக்குங்கள்.
  - **Turing Test:** மேயரின் கூற்றான "மனிதனுக்கும் கணினிக்கும் வித்தியாசம் தெரியவில்லை" என்பதை டியூரிங் சோதனையுடன் நேரடியாகத் தொடர்புபடுத்துங்கள். இயந்திரங்கள் மனிதனைப்

போல் சிந்திப்பதன் உச்சக்கட்ட அளவீடு இதுவே என்று அதன் முக்கியத்துவத்தை எழுதுங்கள்.
    - **PEAS Table:** பரீட்சையில் இதனை ஒரு அட்டவணையாக (Table) எழுதுவது உங்களுக்கு அதிக புள்ளிகளைப் பெற்றுத் தரும்.
        - **P (Performance):** பயண நேரம் குறைதல், எரிபொருள் சேமிப்பு.
        - **E (Environment):** நகர வீதிகள், வாகனங்கள், பாதசாரிகள்.
        - **A (Actuators):** டிராஃபிக் சிக்னல் விளக்குகள், டிஜிட்டல் அறிவிப்புப் பலகைகள்.
        - **S (Sensors):** CCTV கேமராக்கள், வீதிகளில் உள்ள பிரஷர் சென்சார்கள்.

💡 Q2: Basics of Machine Learning (Supermarket Case Study)

- அணுகுமுறை: கொடுக்கப்பட்ட சினாரியோவிற்கு எந்த ML அப்ரோச் பொருத்தமானது எனத் தேர்வு செய்து அதற்கான காரணத்தை (Justification) எழுத வேண்டும்.
- வழிகாட்டல்:
    - i. Customers stopping shopping: இது **Supervised Learning (Classification)**. ஏனெனில், வாடிக்கையாளர் கடற்பாரா (Yes) அல்லது விலகுவாரா (No) என்ற இரண்டு பிரிவுகளில் ஒன்றை லேபிளிடப்பட்ட பழைய தரவுகளை வைத்து மட்டுமே கணிக்க முடியும்.
    - ii. Products purchased together: இது **Unsupervised Learning (Association Rule Mining)**. வாடிக்கையாளர்கள் வாங்கும் பொருட்களுக்கு இடையே உள்ள மறைமுகத் தொடர்புகளைத் (Patterns) தரவுகளில் எந்த லேபிளும் இல்லாமலேயே கண்டறிய இது உதவும்.
    - iii. Delivery Robot Route: இது **Reinforcement Learning**. ரோபோ கிடங்கிற்குள் நகரும்போது சரியான பாதைக்கு ரிவார்டுகளும், தவறான பாதைக்கு பெனால்டிகளும் பெற்று, சோதனைகள் (Trial and Error) மூலம் மிகக் குறுகிய பாதையைக் கற்றுக்கொள்ளும்.
    - **IF-ELSE** வாதம்: மேலாளரின் கருத்தை நீங்கள் மறுக்க வேண்டும் **(Disagree)**. புரோகிராமர் எழுதும் நிலையான விதிகளுக்குப் பதிலாக, ML மாடல்கள் தரவுகளில் இருந்து விதிகளைத் தாங்களாகவே உருவாக்குகின்றன என்பதை உதாரணத்துடன் விளக்குங்கள்.

💡 Q3: Practical - Data Preprocessing & Evaluation Matrix

- அணுகுமுறை: இதில் குறியீடுகள் (No codes) எழுதத் தேவையில்லை. தரவு தயாரிப்பு மற்றும் மாடல் மதிப்பீட்டைப் பற்றியது.
- வழிகாட்டல்:
    - **Problems without Preprocessing:** விடுபட்ட மதிப்புகள் அல்காரிதம்களை முடக்கும், டூப்ளிகேட்கள் Overfitting-ஐ உண்டாக்கும், வெவ்வேறு அளவீடுகள் (Scales) பெரிய எண்களுக்கு மட்டும் மாடல் முக்கியத்துவம் கொடுக்கும்படி செய்துவிடும் என மூன்றையும் விளக்குங்கள்.
    - **Libraries Use: Pandas** (தரவு சுத்தம் செய்தல்), **NumPy** (மேட்ரிக்ஸ் கணிதம் மற்றும் நார்மலைசேஷன்), **Scikit-learn** (மாடல்

ட்ரெய்னிங் மற்றும் எவலுவேஷன்) ஆகிய மூன்றையும் தேர்வு செய்து அவற்றின் பங்களிப்பை விளக்குங்கள்.
    - **Beyond 98% Accuracy:** வெறும் அக்யூரசியை மட்டும் நம்பக்கூடாது. தரவுகளில் உள்ள **Data Imbalance** (ஒரு கிளாஸ் அதிகமாக இருத்தல்), **Precision/Recall** இன் தேவை, புதிய தரவில் மாடல் எவ்வாறு இயங்குகிறது (**Overfitting** சோதனை), மற்றும் மாடல் இயங்குவதற்கான வேகம் (**Latency**) ஆகிய 4 விஷயங்களை ஆராய வேண்டும் என விளக்குங்கள்.

💡 Q4: Two Models (Confusion Matrix Calculations)

- அணுகுமுறை: கணிதக் கணக்கீடுகள் மற்றும் மிகச் சரியான மாடலைத் தேர்வு செய்தல்.
- வழிகாட்டல்:
    - **Model X Calculations:**
        - $\text{Accuracy} = \frac{45 + 32}{45 + 32 + 8 + 15} = \frac{77}{100} = \mathbf{77\%}$
        - $\text{Recall} = \frac{45}{45 + 15} = \frac{45}{60} = \mathbf{75\%}$
    - **Model Y Calculations:**
        - $\text{Accuracy} = \frac{30 + 28}{30 + 28 + 12 + 20} = \frac{58}{100} = \mathbf{58\%}$
        - $\text{Recall} = \frac{30}{30 + 20} = \frac{30}{50} = \mathbf{60\%}$
    - **Suitability (பொருத்தம்): Model X** மிகவும் பொருத்தமானது. ஏனெனில் அதன் ஒட்டுமொத்த துல்லியமும் (77%) மற்றும் ஸ்பேம்களைக் கண்டறியும் திறனும் (Recall 75%) Model Y ஐ விட மிக அதிகமாக உள்ளது.
    - **Different Metrics, Different Conclusions:** வணிகத் தேவைகளைப் பொறுத்து மெட்ரிக்குகள் மாறும். தவறான பாசிட்டிவ்களைத் தவிர்க்க Precision-உம், உண்மையான விடுபடல்களைத் தவிர்க்க Recall-உம் முக்கியம் என்பதால் வெவ்வேறு மெட்ரிக்குகள் வெவ்வேறு முடிவுகளைத் தரும் என விளக்குங்கள்.

💡 Q5: Deep Learning & Generative Models (GANs)

- அணுகுமுறை: நியூரல் நெட்வொர்க் மற்றும் அட்வான்ஸ்டு ஜெனரேட்டிவ் AI பற்றிய கேள்வி.
- வழிகாட்டல்:
    - **Match the blanks / Puzzle:**
        - ReLU $\rightarrow$ E. Prevents negative outputs
        - Epoch $\rightarrow$ C. One complete pass through training data
        - Loss Function $\rightarrow$ A. Measures prediction error
        - Backpropagation $\rightarrow$ B. Updates weights during learning
        - Sigmoid $\rightarrow$ D. Produces outputs between 0 and 1
        - Overfitting $\rightarrow$ F. Performs well on training data but poorly on new data
    - **Activation Functions:** கையெழுத்துப் பிரதிகளை (Handwritten digits) அடையாளம் காணும் போது நேரியல் அல்லாத சிக்கலான வடிவங்கள் (Curves, Edges) இருக்கும். நியூரல் நெட்வொர்க்கிற்கு இந்த சிக்கலான வடிவங்களைப் புரிந்துகொள்ளும் நான்-லீனியர்

(Non-linearity) தன்மையை வழங்க இவை அவசியம் என எழுதுங்கள்.

- **GANs Section:**
  - **Generative vs Classification:** கிளாசிஃபிகேஷன் மாடல் ஒரு படம் என்னவென்று லேபிள்களைக் கணிக்கும்; ஜெனரேட்டிவ் மாடல் முற்றிலும் புதிய தரவை உருவாக்கும்.
  - **Two Components: Generator** (போலிப் படங்களை உருவாக்குபவர்) மற்றும் **Discriminator** (அவை நிஜமா போலியா எனக் கண்டுபிடிக்கும் துப்பறியாளர்) ஆகிய இருவரின் போட்டித் தன்மையை விவரியுங்கள்.
  - **Improvement:** பல சுற்றுகள் பயிற்சி பெறும்போது (Many rounds of training), Discriminator இன் கண்டுபிடிப்புகள் மூலம் Generator தனது தவறுகளைத் திருத்தி, நிஜப் படங்களுக்கு இணையாகத் துல்லியமான படங்களை உருவாக்கக் கற்றுக்கொள்கிறது.