Categorizing Misunderstandings in LLMs:

- Scope of misunderstanding:
    - ~~SM - Description: A short description of what led to an incorrect answer by the LLM. This requires reading the LLM explanation and guessing what went wrong~~
    - SM - Misunderstanding General: These are the general cause of misunderstanding by the LLM. Use one of the three key words given below. In case more than one category applies, separate them with a comma
        - **Wrong Facts/Concept:** Wrong information/hallucinations (when raw facts stated in the answer are wrong) or conceptual misunderstanding (often due to wrong information, but this is something that is implied in the explanation rather than directly stated)
        - **Misinterpreting questions:** Misunderstanding context or implied background in the question. Could mean responding more narrowly or partially (e.g. didn't read the sentence before that gives context). Could mean that the goal of the goal of the question was not understood
        - **Incorrect reasoning/deduction:** This is when the facts in the answer are correct but the conclusion is incorrect. Includes correct answer given in the explanation but wrong choice selected
    - SM - Misunderstanding Reasons: For each category selected in the previous answer for “SM - Misunderstanding General”, provide specific reasons for what caused the LLM to provide an incorrect answer:
        - Wrong Facts/Concept:
            - **Corner Case** - when some information is correct but there is a mistake in a corner case/exception
            - **Out of date information**
            - **Incorrect information/concept** (this can include wrong justifications due to the absence of information e.g. saying any ip address can be of any range) and can include incorrect information/facts
        - Misinterpreting questions:
            - **Quantifier issue:** Question is asking if something is always the case but the answer replies with something that is sometimes the case (and vice versa). This also includes when a question is asking about a typical case but the answer is about a possible case
            - **Direct vs Indirect Causation:** When the question is asking for the cause of an action/phenomenon and the answer is about some cause of the cause of the action. For example: “Question: What causes a SYN+ACK reply? Answer: An application wanting to communicate to another application at a different host”
            - **Misinterpreting a word**: Interpreted a word or background incorrectly (e.g. answering for subnets when the question is talking

about default subnets). Includes ignoring a word or not understanding logical context
        - **Incorrect copying of the question**
        - **No explanation given**
    - Incorrect reasoning/deduction:
        - **Incorrect calculation or counting**
        - **An error related to misinterpreting IP addresses** (also includes incorrect octet mapping e.g. saying last octet has 16 bits)
        - **Contradictory reasoning:** Believing in two contradictory facts/concepts at the same time
        - **Senseless** (e.g. completely random reasoning)
        - **Faulty inference:** A statement does not follow from previous statement
        - **Self-aware but still wrong conclusion**: Explaining how the answer is not valid but still going with it - similar to justifying hallucination. Making things up for no reason
        - **Incorrect Choice:** Deriving the correct answer but still choosing the wrong choice
- ~~SM - Misunderstanding General (Secondary):~~
- ~~SM - Misunderstanding Reasons (Secondary):~~
- ~~Sources~~
    - ~~Source links working: Number of source links that are working. For links, click the link to see if the site is accessible. For books and chapters etc. search online to see if they are real books~~
    - ~~Sources Types: For each source, list the type of source separated by commas.~~ ~~**wikipedia, blog, calculator, RFC, documentation, trusted article, service description, book, slides, forum, research paper, registry**~~
    - ~~Sources Relevant: Number of sources that are relevant to the question~~
- Answer Quality:
    - AQ - Inferrable(0-2): Correct answer can be inferred from the explanation
        - **0** if correct answer not there
        - **1** if the correct answer is there but not obvious (e.g. in case of contradictory answers or poor reasoning - experts should be able to figure it out). If 1 is selected, the concept needs to be correct but there can be calculation mistakes or very basic knowledge missing in the explanation
        - **2** if correct answer is there and is obvious (e.g. not seeing a choice or just selecting the wrong choice - anyone with basic knowledge about semantics should be able to figure the correct answer)
    - AQ - Precise?: If the answer is precise. Mark it as 0 if it is unclear what the answer is from the explanation (throwing stuff at the wall). Hallmark of

an imprecise answer is if there are statements in the explanation that are not necessary?: **0 or 1** (0 if the answer is not precise, 1 otherwise)
        - AQ - Explainable?: Was the answer simply correct, or did it also convey *why* the answer was correct? Did it give an explanation that makes sense? This is regarding the sufficiency of explanations given. e.g. ignoring the unnecessary information, does the explanation sufficiently clarify the answer?: **0 or 1** (0 if the answer is not explanaiable, 1 otherwise)
    - Effects:
        - Effect - Conceptual error in explanation?(0/1): Can the explanation cause conceptual error? (usually occurs when LLM has given a misleading explanation). Has to either be explicit in the explanation or should cause an obvious conceptual error based on the wrong answer
        - ~~Effect - Subtopics: This relates to the kind of effect that reading the incorrect answer can potentially cause. One of **Basic networking** (IP Translation, bgp, subnets etc.)**, Network security** (DDoS, WAN Security etc.)**, Network administration** (Configuring switches, WLAN Configuration etc.)**, Advanced networking** (Ethernet GRE, FlowVisor etc.)~~
    - Correcting Difficulty: Difficulty of correcting the misunderstanding (this is related to the kind of mistake made and not the difficulty of the question)
        - Detection: (1) Quick (rechecking) vs (2) Detailed Reading (might include calculations) vs (3) Would likely not catch it without reading of external document/searching internet
            - Done for only for students
        - ~~Correction: (1) Rechecking (e.g. they would already know the issue and most would be able to correct it from memory), (2) Quick Search (e.g. wikipedia), (3) Reading networking text, (4) Reading documentation/Research Papers, (5) Reading RFC, (6) Reading forums/blogs (e.g. rare questions not immediately found from google), (7) Asking expert (8) Getting into the guts of implementation~~
            - ~~Done for both new students and experts (prof, network op)~~