【문 1】 군집분석(Cluster Analysis)에 관한 다음 설명 중 가장 옳지 않은 것은?

① 군집 내 객체 간의 유사성은 최대화하고, 서로 다른 군집 간의 유사성은 최소화하는 것을 목표로 한다.
② 비계층적 군집화의 대표적인 알고리즘으로는 주어진 개수의 군집으로 객체들을 할당하는 K-means 알고리즘이 있다.
③ 계층적 군집화는 유사한 군집끼리 묶어가는 방식으로, 주요 방법으로는 단일 연결법, 완전 연결법, 와드 연결법 등이 있다.
④ 군집 분석은 사전에 정의된 범주(Category)에 따라 데이터를 분류하는 지도 학습(Supervised Learning)의 대표적인 방법이다.

【문 2】 2024년에 A학술지에 출판된 논문의 수가 300편이고, 이들 논문이 2024년에 900회, 2025년에 600회 인용되었다면 A학술지의 즉시성 지수(Immediacy Index)는 얼마인가?

① 2 ② 3 ③ 4 ④ 5

【문 3】 아래 시소러스 표현을 보고, 이에 관한 설명으로 <보기>에서 옳은 것을 모두 고른 것은?

도서관
  NT 공공도서관
  NT 대학도서관
  RT 사서
  RT 문헌정보학
전자도서관
  USE 디지털도서관
디지털도서관
  UF 전자도서관
  UF 벽 없는 도서관

< 보 기 >
㉠ '전자도서관'은 비디스크립터이며, 색인 시 '디지털도서관'으로 대체하여 사용해야 한다.
㉡ '공공도서관'과 '대학도서관'은 '도서관'과 등가관계에 있다.
㉢ '사서'는 '도서관'과 연관관계에 있다.
㉣ '벽 없는 도서관'이 디스크립터이고 '디지털도서관'이 비디스크립터이다.

① ㉠, ㉢
② ㉠, ㉡, ㉢
③ ㉡, ㉢, ㉣
④ ㉠, ㉡, ㉢, ㉣

【문 4】 다음 중 MeSH에 관한 설명으로 옳은 것을 모두 고른 것은?

㉠ MeSH용어는 개념간의 관계를 수평적인 구조로 배열한 표이다.
㉡ 비통제어휘(uncontrolled vocabulary)이다.
㉢ MEDLINE/PubMed 데이터베이스에서 사용되고 있다.
㉣ MeSH는 Medical Subject Headings의 약자이다.

① ㉠, ㉡
② ㉡, ㉢
③ ㉢, ㉣
④ ㉠, ㉣

【문 5】 메타데이터의 인코딩(Encoding) 구문 표준 및 마크업 언어에 관한 다음 설명 중 가장 옳지 않은 것은?

① HTML: 이식성이 높으나 한정된 태그 구조로 인해, 복잡한 문서 구조나 임의의 레이아웃을 표현하는 데 한계가 있다.
② XML: HTML의 구조적 한계와 SGML의 복잡성을 보완하여, 웹상에서 구조화된 문서를 효율적으로 전송하기 위해 W3C가 제안한 표준 메타언어이다.
③ SGML: 강력한 문서 구조 정의 기능을 가졌으며, 초기부터 웹 환경을 적극 고려하여 개발되어 현재 웹 메타데이터 인코딩의 핵심 표준으로 쓰인다.
④ RDF: 웹 자원의 메타데이터 정의를 위한 개념적 모델로, 데이터를 자원(subject), 속성(predicate), 속성값(object)의 트리플(triple) 구조로 표현한다.

【문 6】 동시출현분석(Co-occurrence Analysis) 및 동시출현 단어분석(Co-word Analysis)에 관한 다음 설명 중 가장 옳지 않은 것은?

① 동시출현분석은 분석 단위에 따라 동시출현 단어분석, 공저자(co-authorship) 분석, 동시분류(co-classification) 분석 등으로 구분할 수 있다.
② 동시출현 단어분석은 두 단어가 동일 문헌에 동시에 출현하는 빈도를 측정하여 단어 간의 연관성을 산출하는 기법이다.
③ 동시출현 단어분석의 주요 가정은 '동시출현 빈도가 낮을수록 두 단어 사이의 개념적 관계가 밀접하다'는 것이다.
④ 동시출현 단어분석의 주요 단계로는 키워드 추출 및 정제, 행렬 작성, 네트워크 분석 등을 활용한 시각화, 지적 구조 해석 등이 포함된다.

【문 7】 정보검색에서 단어의 가중치를 계산할 때 사용하는 TF-IDF 기법에 관한 다음 설명 중 가장 옳은 것은?

① TF는 특정 문헌 집단 전체에서 해당 단어가 출현한 총 횟수를 의미한다.
② IDF는 어떤 단어가 소수의 특정 문헌에만 집중적으로 나타날 때 값이 작아진다.
③ 흔하게 쓰이는 관사나 조사 등은 많은 문헌에 등장하므로 IDF 값이 매우 높게 나타난다.
④ 특정 단어가 한 문헌 내에서 자주 출현하고, 전체 문헌 집단에서는 적은 문헌에만 출현할수록 해당 문헌에서 그 단어의 가중치는 높아진다.

【문 8】 정보격차의 유형에 관한 다음 설명 중 가장 옳지 않은 것은?

① 접근격차는 정보기기와 같은 물리적.물질적 자원에 대한 접근의 기회 격차로, 후속 단계들에 대한 선행요소이자 기반 요건이다.
② 역량격차는 개인 간에 정보기기와 정보를 효과적으로 이용할 수 있는 능력의 차이에서 발생하며, 정보기술에 대한 접근만으로도 정보격차 해소가 가능하다는 것을 의미한다.
③ 질적 이용격차는 이용의 다양성과 활용 수준의 격차를 의미하며, 정보기기를 이용해 얼마나 다양한 온라인서비스를 이용하는가와 얼마나 유용하고 생산적으로 활용하는가와 관련이 있다.
④ 성과격차는 동일한 접근성과 이용능력을 갖추고 있더라도 정보를 어떻게 활용하고 어떤 부가가치를 창출하느냐에 따라 정보혜택에서 차이가 발생할 수 있음을 의미한다.

【문 9】조합색인(coordinate index)에 관한 다음 설명 중 가장 옳지 않은 것은?

① 조합색인이란 두 개 이상의 색인어를 조합하여 문헌의 주제를 만들어 내는 색인법이다.
② 전조합색인은 복합주제를 하나의 색인어로 표현하여 한 번의 검색으로 직접 접근할 수 있도록 한다.
③ 후조합색인은 탐색 이전에 색인작성자가 색인어를 미리 조합해 두기 때문에 이용자가 별도로 개념을 조합할 필요가 없다.
④ 전조합색인의 대표적인 예로는 주제명색인과 분류색인이 있다.

【문10】인공지능기본법에서 설명하고 있는 내용으로 가장 옳지 않은 것은?

① 인공지능윤리: 인간의 존엄성에 대한 존중을 기초로 하여, 국민의 권익과 생명.재산을 보호할 수 있는 안전하고 신뢰할 수 있는 인공지능사회를 구현하기 위하여 인공지능의 개발, 제공 및 이용 등 모든 영역에서 사회구성원이 지켜야 할 윤리적 기준
② 초연결지능정보통신망: 다양한 수준의 자율성과 적응성을 가지고 주어진 목표를 위하여 실제 및 가상환경에 영향을 미치는 예측, 추천, 결정 등의 결과물을 추론하는 인공지능 기반 시스템
③ 인공지능: 학습, 추론, 지각, 판단, 언어의 이해 등 인간이 가진 지적 능력을 전자적 방법으로 구현한 것
④ 생성형 인공지능: 입력한 데이터의 구조와 특성을 모방하여 글, 소리, 그림, 영상, 그 밖의 다양한 결과물을 생성하는 인공지능시스템

【문11】특정 연구자의 학문적 성과와 영향력을 평가하기 위해 널리 사용되는 지표로, "어떤 연구자가 발표한 논문 중 N회 이상 인용된 논문이 N편일 때의 N의 최댓값" 을 의미하는 것은?

① 피인용 지수(Impact Factor)
② h-지수(h-index)
③ 아이겐팩터 지수(Eigenfactor Score)
④ 즉각성 지수(Immediacy Index)

【문12】정보검색 모델 중 '벡터 공간 모델(Vector Space Model)' 에 관한 다음 설명 중 가장 옳지 않은 것은?

① 문헌과 질의를 다차원 공간의 벡터로 표현한다.
② 질의와 문헌 간의 유사도를 코사인 유사도(Cosine Similarity) 등을 통해 계산한다.
③ 문헌 내 단어의 가중치를 계산하기 위해 주로 TF-IDF 기법을 사용한다.
④ 질의 용어가 문헌에 정확히 일치(Exact Match)해야만 검색 결과로 출력되는 엄격한 이진(Binary) 방식을 취한다.

【문13】디지털 도서관과 오픈액세스 환경에서 논문, 저널, 데이터 세트 등의 디지털 객체에 부여되는 영구적인 국제 표준 식별자로, 웹 주소(URL)가 변경되더라도 자원에 안정적으로 접근할 수 있게 해주는 것은?

① DOI (Digital Object Identifier)
② ISBN (International Standard Book Number)
③ ISSN (International Standard Serial Number)
④ ORCID (Open Researcher and Contributor ID)

【문14】메타데이터의 '내용규칙(Content Rules)' 에 관한 다음 설명 중 가장 옳지 않은 것은?

① 내용규칙은 메타데이터 요소의 값이 어떤 방식으로 정보원으로부터 추출되고 표현되는가에 대한 규칙을 제시하여, 메타데이터 작성의 정확성 및 일관성을 확보하는 역할을 한다.
② '요소의 값을 선정하는 방법'의 예로는, 특정 정보자원에 본표제와 축약표제 등 여러 서명이 존재할 때 그 중 어떤 서명을 본표제로 채택할지 기준을 정하는 것을 들 수 있다.
③ '요소의 값을 표현하는 방식'은 메타데이터 작성 시 값을 자연어로 자유롭게 기술할 것인지, 아니면 통제어휘집에서 추출한 통제어를 사용할 것인지에 대한 기준을 정하는 것을 의미한다.
④ 더블린 코어(Dublin Core)에서 날짜를 표기할 때 'W3CDTF' 형식을 지정하여 'YYYY-MM-DD'와 같이 특정 형식으로 입력하도록 표준화하는 것은 '요소의 값을 표현하는 방식'에 해당한다.

【문15】구글의 페이지랭크(PageRank Algorithm)에 관한 설명으로 옳은 것을 모두 고른 것은?

㉠ 웹 페이지의 상대적 중요성을 측정하기 위해 사용된다.
㉡ 중요한 페이지일수록 실제 페이지랭크 값은 높게 된다.
㉢ 어떤 질의를 처리하고 있는지에 따라 웹 페이지는 다른 페이지랭크 값을 갖는다.
㉣ 페이지랭크에서 모든 페이지랭크 값의 합은 0이 된다.

① ㉠, ㉡
② ㉡, ㉢
③ ㉢, ㉣
④ ㉠, ㉣

【문16】다음 파일 포맷 중 장기 보존포맷으로 가장 적합한 포맷으로만 옳게 짝지어진 것은?

① MPEG - GIF
② XML - HWP
③ GIF - XLS/XLSX
④ PDF/A - TIFF

【문17】다음 중 특정 커뮤니티(분야)와 해당 분야에서 주로 사용되는 메타데이터 표준(Metadata Standard)의 연결이 가장 옳지 않은 것은?

① 전자도서관: MODS (Metadata Object Description Schema)
② 출판계: ONIX (ONline Information eXchange)
③ 예술작품: LOM (Learning Object Metadata)
④ 연방정부 정보: GILS (Government Information Locator Service)

【문18】오픈액세스(OA) 리포지터리에 관한 다음 설명 중 가장 옳지 않은 것은?

① 기관 리포지터리용 플랫폼으로는 Dspace와 Eprints 등이 있다.
② OpenDOAR은 OA 리포지터리를 검색할 수 있는 목록을 구축하여 제공하는 서비스이다.
③ 주제 리포지터리는 특정 학문 분야나 주제의 연구성과물이나 데이터를 보관하는 온라인 아카이브이다.
④ 기관 리포지터리는 저자 셀프 아카이빙을 통해 데이터가 축적된다.

【문19】 ISO/IEC에서 채택한 국제표준인 토픽 맵(Topic Maps)의 핵심 구성요소로 가장 옳게 묶인 것은?

① 주제(topic) - 연관(association) - 발현(occurrence)
② 자원(resources) - 속성(property) - 값(value)
③ 개체(individual) - 속성(property) - 데이터 타입(type)
④ 주어(subject) - 술어(predicate) - 목적어(object)

【문20】 부다페스트 선언(BOAI, 2002)에서 제시한 오픈액세스(OA)의 정의에 포함된 내용으로 가장 옳지 않은 것은?

① 이용자 누구나 학술지 논문에 무료로 접근할 수 있어야 한다.
② 법률적.기술적 장벽 없이 논문을 이용할 수 있어야 한다.
③ 저작에 대한 통제권은 출판사가 가지며, 인용 시 출처를 밝혀야 한다.
④ 특별한 허가절차 없이 최대한 재사용 가능한 상태여야 한다.

【문21】 Creative Commons(CC) 라이선스의 기본 조건 요소 중 모든 CC 라이선스에 공통으로 포함되는 필수 조건은?

① ND(변경금지)
② SA(동일조건변경허락)
③ BY(저작자표시)
④ NC(비영리)

【문22】 저작권법에 기술된 저작물 보호에 관한 다음 설명 중 가장 옳지 않은 것은?

① 편집저작물은 독자적인 저작물로서 보호된다.
② 원저작물을 번역.편곡.변형.각색.영상제작 그 밖의 방법으로 작성한 창작물은 독자적인 저작물로서 보호받지 못한다.
③ 2차적저작물의 보호는 그 원저작물의 저작자의 권리에 영향을 미치지 아니한다.
④ 사실의 전달에 불과한 시사보도는 저작물로서 법에 의해 보호받지 못한다.

【문23】 검색성능 평가 척도 중 정확률(precision)에 관한 다음 설명 중 가장 옳은 것은?

① 검색시스템이 부적합문헌을 검색해 내지 않는 능력이다.
② 검색되지 않은 적합문헌의 비율이다.
③ 검색된 부적합문헌의 비율이다.
④ 정확률 40%는 검색된 문헌 10개 가운데 6개가 적합문헌임을 의미한다.

【문24】 다음 중 더블린 코어(Dublin Core) 메타데이터의 15개 기본 요소에 해당하지 않는 것은?

① Title(표제)
② Creator(저작자)
③ Affiliation(소속기관)
④ Format(형식)

【문25】 베르지히와 네벨링(Wersig & Neveling, 1975)이 분류한 정보학에 대한 견해 중, 사서나 분류이론 및 시스템 설계에 관심을 가지는 실무 중심의 부류에 해당하는 것은?

① 사상 지향적 견해(phenomenon-oriented view)
② 방법 지향적 견해(means-oriented view)
③ 기술 지향적 견해(technical-oriented view)
④ 목적 지향적 견해(purpose-oriented view)