

# 빅데이터분석 기사 국가기술자격 개요 및 출제 경향 분석

장희선<sup>o</sup>, 송지영<sup>\*</sup>

<sup>o</sup>평택대학교 융합소프트웨어학과,

<sup>\*</sup>한국폴리텍대학 반도체융합캠퍼스 반도체융합SW과

e-mail: hsjang@ptu.ac.kr<sup>o</sup>, jysong1@kopo.ac.kr<sup>\*</sup>

## An Introduction and Trend Analysis in Questions of Engineer Big Data Analyst

Hee-Seon Jang<sup>o</sup>, Ji Young Song<sup>\*</sup>

<sup>o</sup>Dept. of Convergence Software, Pyeongtaek University,

<sup>\*</sup>Dept. of Semiconductor Convergence Software Campus of Korea Polytechnic

### ● 요약 ●

본 논문에서는 과학기술정보통신부와 통계청에서 주관하고 한국산업인력공단에서 시행(한국데이터산업진흥원 위탁)하는 「빅데이터분석기사」에 대한 필기 및 실기 시험의 내용을 설명하고 지금까지 2회에 걸쳐 시행된 시험에 대한 문제점과 이에 대한 해결방안을 제시하였다. 2021년 처음 시행된 국가기술자격으로써 기존 자격증과의 차별성, 난이도 조정, 수험생들의 각종 민원 발생 등의 문제를 해결하기 위한 체계적인 시스템 마련이 요구되며, 향후 데이터 과학자들에 대한 수요 급증에 대비하기 위해 빅데이터분석 실무 능력을 평가하기 위한 바람직한 제도와 정책이 병행되어야 한다.

**키워드:** 빅데이터분석기사(big data analyst), 데이터과학자(data scientist), 인공지능(artificial intelligence)

### I. Introduction

Klaus Schwab이 4차 산업혁명이라는 키워드를 던진 이후, 최근 인공지능(AI; Artificial Intelligence) 서비스를 제공하기 위해 반드시 요구되는 빅데이터분석 분야에 대한 인력 수요가 급증하고 있다 [1-4,7,8]. 이에 과학기술정보통신부와 통계청에서는 2021년부터 「빅데이터분석기사」에 대한 국가기술자격 시험을 통해 인력난을 해소하고자 노력하고 있다[1,5]. 빅데이터분석기사는 국가기술자격법에 의해 2021년부터 시행되었으며 “빅데이터 이해를 기반으로 빅데이터 분석 기획, 빅데이터 수집저장처리, 빅데이터 분석 및 시각화를 수행하는 실무자”로 정의하고 있다[6].

한국데이터산업진흥원에서 발간한 데이터산업 백서[5]에 따르면 향후 5년 내 데이터 산업을 포함한 전 산업에서 빅데이터 직무와 관련된 필요 인력은 2배(데이터 및 전 산업) 이상 필요(그림 1)로 하고 있다. 특히, [그림 2]에서와 같이 데이터 산업 직무별 인력 부족률에서 보면, 데이터과학자(34.2%), 데이터분석가(14.8%), 데이터개발자(13.5%) 순으로 인력이 필요함을 알 수 있다.

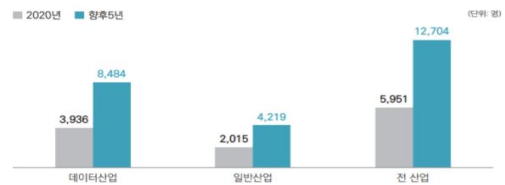


Fig. 1. 빅데이터 직무 필요 인력[5]

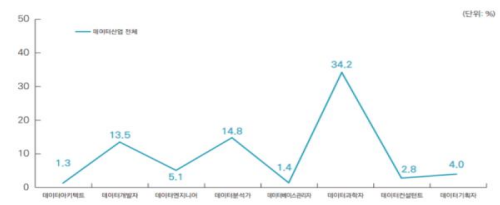


Fig. 2. 데이터 직무별 인력 부족률[5]

## II. Engineer Big Data Analyst

빅데이터분석기사 국가기술자격의 주무부처는 과학기술정보통신부와 통계청이며, 한국산업인력공단에서 관리하고, 한국데이터산업진흥원에서 위탁시행[1,5]하고 있다. 시험 과목과 합격 기준은 다음과 같다.

① 시행처: 한국데이터산업진흥원 ② 시험과목 - 필기: 1. 빅데이터 분석 기획 2. 빅데이터 탐색 3. 빅데이터 모델링 4. 빅데이터 결과해석 - 실기: 빅데이터 분석 실무 ③ 점정방법 - 필기: 객관식, 과목당 20문항 - 실기: 복합형(문답형, 작업형) ④ 합격기준 - 필기: 100점을 만점으로 하여 과목당 40점 이상, 전과목 평균 60점 이상 - 실기: 100점을 만점으로 하여 60점 이상
---

필기는 4과목(분석기획, 탐색, 모델링, 결과해석), 실기는 데이터분석 실무로써 필답형(10문제(30점))과 작업형(1유형 3문제(30점), 2유형 1문제(40점))으로 구성되어 R 또는 Python을 이용한 실무 능력(데이터 수집, 전처리, 모형 구축, 모형 평가 작업)을 테스트 한다. 그리고 한국데이터산업진흥원에서는 아래와 같이 직무를 정의하고 그 필요성을 강조하고 있다.

● 빅데이터분석기사 정의 빅데이터 이해를 기반으로 빅데이터 분석 기획, 빅데이터 수집·저장·처리, 빅데이터 분석 및 시각화를 수행하는 실무자를 말한다. ● 빅데이터분석기사의 필요성 전 세계적으로 빅데이터가 미래성장동력으로 인식되며, 각국 정부에서는 관련 기업투자를 끌어내는 등 국가기업의 주요 전략분야로 부상하고 있다. 국가와 기업의 경쟁력 확보를 위해 빅데이터 분석 전문가의 수요는 증가하고 있으나, 수요 대비 공급 부족으로 인력 확보에 어려움이 높은 실정이다. 이에 정부차원에서 빅데이터 분석 전문가 양성과 함께 체계적으로 역량을 검증할 수 있는 국가기술자격 수요가 높은 편이다.
---

지금까지 3차(1회 시험은 코로나로 취소)에 걸친 시험결과 필기 합격률은 40%, 실기 합격률은 약 60% 정도를 보이고 있다[1,5].

## III. Trend Analysis in Questions

빅데이터분석기사 필기 문제의 출제 유형을 분석하면 <표 1>과 같다.

Table 1. 빅데이터분석기사 출제 경향[4]

<b>I. 빅데이터 분석 기획</b> 빅데이터 특징, 진단분석, 데이터 과학자 소양, 데이터분석 조직 성숙 단계, 데이터 거버넌스, ETL, 인공지능 학습, 머신러닝, 딥러닝, CNN, 개인정보 처리, 재현·마이 데이터, 데이터분석 단계, 이상값, 데이터 수집, GDPR, 데이터 비식별화, 데이터 품질 점검, 데이터 스케일링, NoSQL, 데이터웨어하우스, Hadoop Ecosystem	<b>II. 빅데이터 탐색</b> 박스 플롯, 데이터 정제, 차원의 저주, 회귀분석, 차원 축소, 주성분 분석, 공분산 행렬, 고유값, 고유벡터, 파생변수, Box-Cox, 불균형 데이터 처리, EDA, 상관계수, 기술통계량, 전수조사, 표본조사, 표본추출 방법, 확률분포(포아송, t-분포, 정규분포), 중심극한정리, 점추정, 신뢰구간 추정, 가설검정, 유의수준, 유의확률
<b>III. 빅데이터 모델링</b> 회귀분석 모형, 모델링 절차, 다중공선성, ANOVA, 로지스틱 회귀분석, 인공지능경량 하이퍼 파라미터, SVM, 커널함수, RBF, k-means, MDS, 시계열 분석, 베이저안 확률, CNN feature map, DBN, 활성화 함수, softmax 함수, 시그모이드 함수, 경사 하강법, 그래프 분석, 앙상블 분석, 랜덤 포레스트, 비모수검정(부호, Wilcoxon, Friedman)	<b>IV. 빅데이터 결과 해석</b> 최대 우도추정법, 혼동행렬, ROC, AUC, MAPE, 모형 진단 절차, 잔차, 다원 분산분석, Hold-out, 교차검증, 적합도 검정, 편향과 분산, 카이제곱 검정 통계량, 과대적합 방지 방법, 시계열 자료 성분, 연관성 분석, 데이터 분석 활용, 분석결과 시각화, 관계 시각화, 산점도, 거미줄차트, 인포그래픽, Cartogram, 분석결과 활용

실기는 단답형(10문제), 작업형 1유형(데이터처리), 작업형 2유형(모형 구축 및 평가)으로 구분하여 주어진 데이터셋을 이용한 데이터 처리, 모형 구축 및 평가(ROC, AUC 등) 업무의 수행(R 또는 Python) 능력을 테스트 한다.

지금까지 2회에 걸쳐 시험을 보면서 ① 문제 오류, ② 난이도 조정 실패, ③ 기존 기사 자격증과의 차별성 미흡, ④ 실기 시험에 대한 객관적 실무능력 평가의 한계, ⑤ 민원에 대한 즉각적인 대응 미흡 등의 문제점이 대두되었다. 향후 이러한 문제점을 고려하여 바람직한 빅데이터분석 실무능력 평가를 위한 지침이 되어야 할 것이다.

## IV. Conclusions

과학기술정보통신부와 통계청에서 주관하는 빅데이터분석기사에 대한 필기실기 시험의 개요를 설명하고 지금까지 시행된 문제 유형을 분석하였다. 데이터과학자에 대한 수요의 증가와 함께 자격증에 대한 중요성이 점점 더 증가할 것으로 예상된다. 향후 실생활에 도움이 되는 빅데이터분석 실무 능력을 평가하기 위한 문제의 실효성 제고 방안과 체계적인 관리를 위한 시스템이 조속히 마련되어야 한다.

## ACKNOWLEDGEMENT

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.2021R1F1A1049933).

## REFERENCES

- [1] Data License Test, <https://www.dataq.or.kr/www/main.do>, 2021.
- [2] D.S Kim, Welcome to the AI World, Life and Power Press, 2021.
- [3] Human Resources Development Service of Korea, <https://www.hrdkorea.or.kr>, 2021.
- [4] HJ Jeong and HS Jang, "Engineer Big Data Analyst", 2022.
- [5] Korea Data Industry Agency, <https://www.kdata.or.kr>, 2021.
- [6] National Competency Standards, <https://ncs.go.kr/index.do>, 2021.
- [7] Michael Negnevsky, Artificial Intelligence, IT Cook Book, Hanbit, 2021.
- [8] Satosy DD, Introduction to Artificial Intelligence, Hanbit, 2021.