

인공지능 기술을 활용한 ‘가짜뉴스 찾기’ 챌린지를 통한 국내 R&D 지원 시스템의 방향성에 대한 연구

A study for the Korean National R&D Policy through 'Fake News Finding' in artificial intelligence Challenge

전 광 호, 하 선 우
정보통신기술진흥센터

Kwang-ho Chun, Sun-woo Ha
IITP

요약

우리 정부는 신기술인 지능정보기술의 효과 실현을 극대화하고, 선진 기술 확보를 가속화하기 위해 기존 추격형을 넘어선 선도형 R&D 지원체계가 필수적인 상황을 인지하고 있다. 특히, 미국 등 선진국에서는 신기술 분야에 대한 효과 극대화 및 R&D 촉진을 위해 도전·경쟁형 R&D 지원체계를 활용한 R&D 추진을 활발히 지원하고 있다. 세계적으로 첨단미개척 인공지능 분야의 연구 진작을 위해 우리정부는 지난 '17년부터 도전형·경쟁형·개방형 R&D 지원 체계인 '인공지능 R&D 챌린지' 도입하여 지원하고 있다. 이에 본 논문에서는 인공지능 R&D 챌린지의 가짜뉴스 찾기를 통해 국내 R&D의 지원방향에 대해 검토해 보았다.

I. 서론

1. 인공지능 R&D 챌린지

우리 정부는 '17년 상반기부터 세계적으로도 첨단미개척 분야인 인공지능 분야 연구 진작을 위해 도전형·경쟁형·개방형 R&D 지원 체계인 'AI R&D 챌린지' 도입하여 지원하고 있다. 높은 경제·사회적 효과가 기대되나 기술 난이도가 높아 해결되지 않은 과제를 발굴하여 관련 R&D 대회 개최 계획을 공고하고, 대회일까지 연구자들이 자력으로 기술 개발에 도전토록하고, 일정 수준 이상의 기술 개발에 성공한 참가자(들)에게 후속 연구비를 지원하는 챌린지 방식의 새로운 R&D 방식을 추진하고 있다. 일정 수준 이상의 성능을 내는 기술 개발에 성공한 참가자가 복수인 경우 복수 연구자에 초기 연구비를 지원하되, 연차 평가를 거쳐 단일 우수 연구기관에 연구비 지원을 집중하는 경쟁형 R&D 도 같이 추진하고 있다.

참가 대상은 대학기업연구소일반 국민 등 제한 없이 모집하며, 국제 공동 연구팀 및 해외 연구기관(국내 법인 조직 소유 필수)도 참여도 허용하여 경연방식의 선형 R&D를 추진하는 것이다. 운영 절차는 과제 공고 후 참가자가 대회 출전을 위한 R&D를 자체 추진하고, 정부가 전문기관의 도움을 통해 대회 통해 평가하여 후속 지원 대상을 결정하는 방식이다. 사업 특징은 경제·사회적 효과가 높으나, 기술 난이도로 인해 해결되지 않은 과제를 발굴하여 R&D이다. 문제 발굴공고는 정부·민간의 수요 공모 과정을 거쳐 후보 문제 pool을 구성, 산학연 전문가로 구성된 문제발굴위원회를 통해 심의한다. 문제 예시로는 가짜뉴스 판별, 고령자 개호, 영유아 보호, 재난 구조, 미세 먼지 대응 등이 초반에 언급되었으며, 최종 1차년도 문제는 가짜뉴스 찾기로 정해졌다. 이에 대한 설

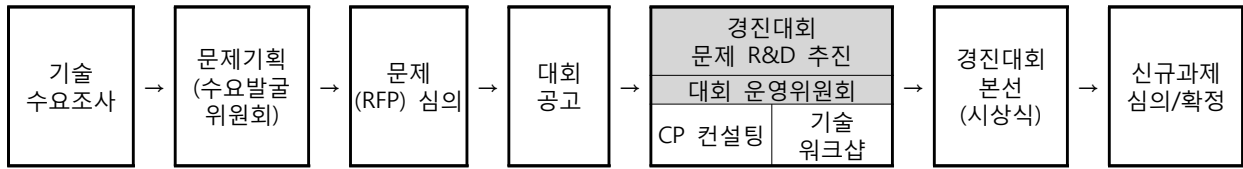
명은 다음 장에서 진행한다. 초기 연구는 대회 참가자는 주어진 기간 동안 완전 자율 방식으로 지능정보기술을 활용한 문제 해결책을 초기 연구를 진행하고, 참가자에 충분한 준비연구기간(1년 이상)을 부여하여 창의력기술력을 바탕으로 문제 해결에 도전토록 배려할 예정이다. 대회 개최는 과제 해결을 위한 연구성과를 공개대회를 통해 경쟁을 하며, 최종 연구비 지급 및 후속 연구는 초기 연구 목표를 달성한 연구자에게 연구 성과의 고도화를 위한 후속 R&D 자금을 연계 지원할 예정이다. 연구 목표를 달성한 모든 팀에게 1년 간 후속 지원금 제공하되 대회 실적에 따라 지원금액 차등지원할 계획이다. 단, 연차별 후속 R&D 자금 지원은 경쟁형 R&D 방식을 적용하여 연차별로 R&D 수행 기관수를 축소하되 우수 팀에 지원금을 집중할 예정이다.

II. 본론

1. 가짜뉴스 찾기

‘속이는 것을 목적으로 만들어진 이야기’로 정의되는 가짜 뉴스는 뉴스 미디어 업계뿐 아니라 혼란으로 야기되는 사회적 비용 낭비를 방지하기 위해서도 관심을 가지고 있는 주제이다. 특히, 국내외에서 SNS를 통한 가짜 뉴스의 확산으로 피해가 발생하는 것을 막기 위해서는 뉴스의 사실 여부를 자동 분석하는 연구에 대한 관심이 높아지고 있다. '17년 인공지능 R&D 대회는 주어진 뉴스 기사를 분석하여 사실여부를 자동으로 판단하는 인공지능 SW개발을 최종 목적으로 한다.

가짜 뉴스 찾기는 다음과 같이 크게 3가지 연구를 통해 접근이 가능하다. ① 뉴스 컨텐츠(제목/내용)의 정합성/불일치성을 통한 검출, ② 뉴스가 사용자들 사이에서



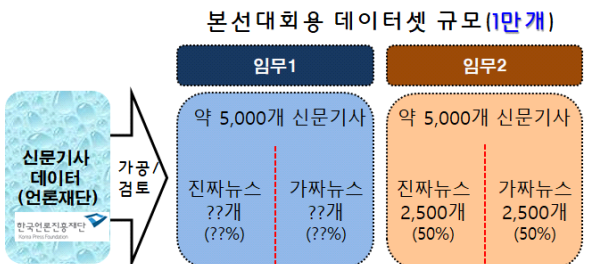
▶▶ 그림 1. 인공지능 R&D 챌린지 경진대회 추진 절차

전파되는 패턴 및 클라우드 기반 검출, ③ 뉴스토픽에 대한 지식 베이스 구축 및 추론을 통한 불일치 검출 등이다. 위의 연구는 궁극적인 가짜 뉴스 찾기에 독립적 또는 통합적으로 활용되어 효과적인 성능을 가져올 수 있다. 본 인공지능 R&D 대회에서는 제한된 시간 내 실현가능성 확인, 난이도 조절의 용이성, 다양한 주제에 적용가능성 등을 고려하여, 가짜 뉴스 찾기의 첫 단계인 ‘뉴스 키텐츠의 정합성/불일치성을 통한 검출’을 문제로 선정하였다.

2. 인공지능 R&D 챌린지 진행

정부는 챌린지 기획을 위해 우선적으로 인공지능 R&D 챌린지 문제발굴위원회를 구성하여 운영하였다. 문제발굴위원회의 검토회의와 문제해결 수요조사(민간 기업·정부기관 대상)를 통해 문제은행(12개) 구성, 온라인 호응도조사 실시 및 평가하여 최종 문제를 가짜뉴스로 결정하였다.[1]

2017 인공지능 R&D 챌린지 과제는 가짜뉴스 검출을 위한 기초·요소기술 개발로 ① 뉴스(기사) 제목과 내용의 정합성 판별, ② 뉴스(기사) 본문 중 맥락에 관계없는 내용 포함여부 검출 등 두가지 미션으로 진행하였다. 평가 기준은 가짜뉴스 판별 솔루션의 뉴스(기사)에 대한 ① 제목·내용 정합성 판별 정확도, ② 맥락에 알맞지 않은 내용 검출 정확도로 판별하였다. 챌린지에서 가장 중요한 ‘가짜뉴스 찾기’ 미션에 필요한 ‘데이터셋’을 챌린저들에게 제공하기 위해 원 소스 기사(디지털 신문기사) 구입, 가공 및 제공하였다.



▶▶ 그림 2. 인공지능 R&D 챌린지 데이터셋 규모

인터넷신문사인 (주)위***의 데이터(디지털 신문기사) 1만건 제공협조를 통해 가공하여 챌린저들에게 제공하여 샘플데이터셋을 제공하였다. 본선의 실질 테스트 데이터셋을 준비하기 위해 IITP가 한국언론진흥재단을 통해 10여개 신문사들의 데이터(디지털 신문기사)를 일괄 구입/제공받아 가공, 본선대회에 사용하였다. 데이터 규모는 약 30만건에 달했다. 1차 워크샵을 통해 배포된 샘플데이터셋 규모 현황은 제공기사 총 700개 내외(미션당 350개 내외), 기사비용 50%(진짜뉴스) : 50%(가짜뉴스)로 챌

린저들에게 제공하였다. 참가팀 71개 팀을 대상으로 예선 및 본선을 통해 공개되는 데이터(1만건 내외)에 대해 참가팀이 자체적으로 실행 후, 결과물을 온라인으로 제출하여 최종 경연을 진행하였다.[3]

3. 국가 R&D의 방향성

동 인공지능 R&D 챌린지는 통해 국가 R&D의 방향성을 제시하고자 한다. 첫째, 정부는 고위험·고수익 (high-risk, high-return) 성격의 과제를 과감히 발굴·후속지원을 제시하여 연구자에 도전 동기 부여하는 도전형 R&D를 추진하여야 한다. 연구자는 실패위험성을 감수하고 문제 해결 가능성 입증을 위한 선행연구를 자율적으로 추진, 후속 연구과제 수행에 도전하여야 한다. 둘째, 복수의 연구자가 다양한 방식으로 문제 해결에 도전하도록 하여 문제 해결의 가능성을 제고하는 경쟁형 R&D를 추진하여야 한다. 성과 경쟁을 통해 우수 참가자를 선발하고 우수자에 지원을 집중하여 ‘경쟁과 집중’ 원리에 따른 R&D 투자 효율 극대화하여야 한다. 첨단미개척 분야의 도전연구 진작을 위한 경쟁형 R&D를 확대하여야 한다. 마지막으로 참가 자격 완화, 접수 기간 확대, 지원 서식 간소화 등 참가자 모집 단계부터 개방성을 확대하여 연구자 다양성 확보하는 개방형 R&D를 추진하여야 한다. 대회 참여를 위한 초기 연구가 완전 자율임은 물론, 후속 R&D 과제의 기획도 연구자가 주도도록 하여 과제 기획의 개방성을 극대화하여야 한다.

III. 결론

본 논문에서는 '17년 인공지능 R&D 챌린지의 ‘가짜 뉴스 찾기’ 대회를 통해 국가 R&D의 향후 방향성에 대해 언급하였다. 도전형, 경쟁형 및 개방형 R&D에 대한 의견을 제시하였다. 결국, 우리나라는 한정된 자원 바탕에서 글로벌 진출을 위한 기반을 다지기 위해서는 국가 R&D는 반드시 전세계적인 트렌드를 따라가야 할 것이며, 선제적인 지원을 통해 우리만의 핵심 기술을 가져야 할 것이다.

■ 참고 문헌 ■

[1] IITP “인공지능 R&D 챌린지 추진 계획(안)”, '17. 4월
 [2] 미래창조과학부 공고 제2017-0298호 “2017년도 인공지능 R&D 챌린지 대회 공고”, '17. 6월
 [3] IITP “2017 인공지능 R&D 챌린지 관련 우승팀 선발을 위한 본선대회”, '17. 12월