

2017 (사)ICT플랫폼학회 하계학술대회 및 (사)ICT대연합 지능정보행정 토론회 논문집

ICTPS



일시 2017년 6월 16일(금) 오후 1시~17일(토) 11시

장소 백석대학교



- 주 최 : (사)ICT플랫폼학회, (사)ICT대연합, 동국대 링크사업단
- 주 관 : (사)ICT플랫폼학회
- 후 원 : 백석대학교, (주)세림티에스지, (주)시스원, (주)콤텍정보통신, (주)세림소프트, (주)굿모닝아이텍, (주)대유플러스, (사)아시아행정학회, (주)티맥스소프트

도메인 지식으로부터 기반 온톨로지 구축을 위한 방법론	75
	안정현, 박용범(단국대학교)

SESSION 2

4채널 Thermocouple를 이용한 온도 센싱 시스템 개발	83
	권민수, 유진호, 강윤희(백석대학교)
희토류 첨가 광섬유와 장주기 광섬유 격자쌍 기반의 전광식 가변형 광감쇠기에 관한 연구	86
	김윤현(한국광기술원), 한원택(광주과학기술원)
모노폴 도파관 여기를 갖는 코니컬 혼 안테나 설계	90
	박동희(한국교통대학교)
Logistic Classification을 통한 태양광 발전시스템의 고장여부 판단방법 구현	94
	박지훈, 손현승(홍익대학교), 서채연(선문대학교), 김영철(홍익대학교)
휴양림 및 무선 음영지역을 위한 LoRa기반 네트워크 프레임워크 제안	98
	박필원(고려대학교), 고대식(목원대학교)
1-Wire로 구성된 DS1820센서를 이용한 온도센싱 시스템 구성	103
	성명훈, 유진호, 강윤희(백석대학교)
사물인터넷을 이용한 반도체 증착 공정의 부하최소화 조절 방법과 설계	106
	조성익, 김정호(한밭대학교)
28GbpsX4채널 고속 광모듈용 세라믹 패키지 개발	110
	정수용, 최광성, 임권섭(한국전자통신연구원), 남동우(주)메탈라이프, 이종진(한국전자통신연구원)

SESSION 3

중소형 의료기관 보안관리체계 설계를 위한 의료분야 특화 보안 요소 도출	117
	김자원(중앙대학교), 김양훈(극동대학교), 장항배(중앙대학교)
저전력 블루투스를 이용한 유아용 체온 측정기	120
	우태훈(금오공과대학교), 조화주((주)아이티로그), 이용환(금오공과대학교)
아두이노를 사용한 난청인을 위한 스마트 팔찌	123
	윤여진, 김석훈(순천향대학교)

Logistic Classification을 통한 태양광 발전시스템의 고장여부 판단방법 구현

Failure Judgement Method of Photovoltaic System through Logistic Classification

박지훈*, 손현승*, 서채연**, 김영철***

Jihoon Park*, Hyun Seung Son, Chae Yun Seo** and R. Young Chul Kim***

요 약

기존의 태양광 모니터링은 후처리 기반의 시스템들이다. 최대 에너지 효율을 위해 태양광 발전소의 고장 예측이 매우 중요하다. 기존은 데이터의 빅 데이터화와 학습 및 고장 판단하는 것이다. 결국 후처리기반이다. 본 논문은 선처리를 위해 고장 예측 방법을 구현했다. 이는 효율적인 에너지 구축을 위해 Logistic Classification을 오류예측 방법을 적용한다. 또한, Logistic Regression을 이용하여 미래의 데이터가 고장데이터일 확률을 수치화할 수 있다. 이를 통해 태양광발전 시스템의 고장횟수가 줄어들어 태양광발전소가 지속적인 운영이 기대된다.

Key words : photovoltaic system, machine learning, tensorflow, classification

1. 서 론

2015년 미국에서 신규로 증설된 태양광 발전 설비용량은 7,260MW로 전년 대비 16% 증가하여 사상 최고 기록을 달성하였다. 2015년 말에는 총 25.6GW에 이르는 것으로 집계되었다. 게다가 전체 신규 발전 설비용량 증설 중 29.5%가 태양광 발전에 의한 것이었다[1]. 세계 신재생에너지 시장은 2020년까지 약 1조 달러에 육박할 것으로 예상된다[2]. 전 세계 평균 태양광 발전 시스템의 가격은 시간당 1MW를 생산하는 기준(MWh)으로 2012년에 비해 절반가량인 99달러로 떨어졌다. 신재생에너지의 보급이 확산되면서 태양광 셀 효율

이 향상되는 등 관련 기술이 발전하였다[3].

하지만 대한민국은 지리적인 여건상 발전 단가를 낮추기가 어려운 편이다. 대한민국은 중남미, 아프리카 등에 비해서 절반가량으로 발전시간이 부족하다. 상대적으로 한정된 공간과 짧은 발전시간으로 최대한의 효율을 내야한다. 만약 태양광발전 시스템이 고장 난다면 태양광 발전소는 큰 타격을 입게 될 것이다. 신재생 에너지와 관련하여 태양에너지에 관한 연구는 활발히 이루어지고 있지만 태양광발전시설에 대한 연구는 미흡한 실정이다[4].

기존연구는 총 3개 년도로 1차로 이중의 인버터간의 모니터링 시스템 구축을 하였다. 2차로 지열 발전소와의 모니터링 시스템을 통합 구축하였다. 3차

* 홍익대학교 소프트웨어공학연구소(pjh@selab.hongik.ac.kr)

** 선문대학교 IT교육학부 강의전담교수(jyun@selab.hongik.ac.kr)

*** 교신저자: 홍익대학교 소프트웨어공학연구소(bob@hongik.ac.kr)

· 제1저자 (First Author) : 박지훈

로 머신러닝을 이용한 고장 예측이다. 본 논문에서는 머신러닝을 통해 태양광 발전 시스템의 고장을 예방하기 위한 방법을 제안한다. 기존에는 후처리를 통한 고장 데이터 분석뿐이었다. 머신러닝 기법 중 한가지인 Logistic Classification을 이용하여 기존의 태양광 발전 모니터링 데이터[2, 5, 6]를 통해 고장 데이터를 확인할 수 있다. 고장 데이터를 확인함으로써 태양광 발전 시스템의 고장여부를 예측할 수 있다.

본 논문은 다음과 같은 순서로 구성되어 있다. 2장에서는 관련 연구로 태양광발전 모니터링 시스템, 텐서플로에 관한 내용에 대해 설명한다. 3장에서는 머신러닝 기법을 적용하여 태양광발전 시스템의 고장을 예측하는 방법에 대해 설명한다. 4장에서는 결론 및 향후 연구에 대해 설명한다.

II. 관련 연구

1. 태양광발전 모니터링 시스템

태양광발전 시스템은 태양광 패널을 통해 얻은 전기 에너지를 자체적으로 사용하거나 발전소로 보낸다. 태양광발전 모니터링 시스템[2]은 이러한 태양광발전 에너지의 수치뿐만 아니라 인버터를 통해 전력, 온도, 경사 센서와 같은 정보도 확인할 수 있다. 그림 1은 태양광발전 모니터링 시스템의 구성도이다. 그림 1을 보면 n개의 로컬 서버가 존재한다. 로컬 서버는 로컬 태양광발전 에너지 생성을 관리한다. RS232 통신을 사용하여 Junction Box들과 인버터들에게 생성된 모니터링 데이터를 전송한다. 이를 통합하는 서버는 모든 데이터

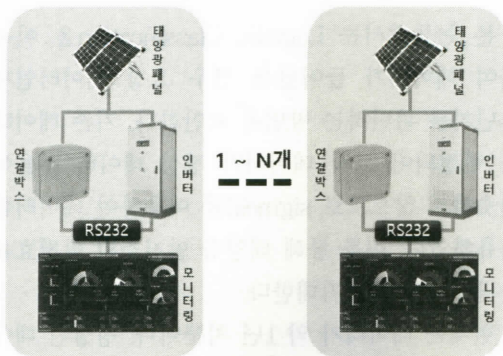


그림 1 태양광발전 모니터링 시스템 구성도

를 관리 할 수 있는 TCP/IP 통신을 사용하여 로

컬 서버에서 변환된 데이터를 수집한다. 서버는 모든 데이터를 데이터베이스에 저장하기 때문에 분석이 필요한 경우 서버는 빅 데이터 시스템을 통해 데이터를 분석할 수 있다. 빅 데이터 시스템은 신재생 에너지의 실시간 예측, 통합 제어, 최적 상태 및 운영, 통합 모니터링과 같은 서비스를 제공할 수 있다. 소프트웨어 플랫폼은 프레임 워크, RS232 통신, Netty 기반의 TCP/IP 통신, Hadoop으로 구성되어 있다.

2. 텐서플로우

텐서플로우(Tensorflow)는 구글 제품에 사용되는 머신러닝을 위한 오픈소스 소프트웨어 라이브러리이다. 구글내 연구와 제품개발을 위한 목적으로 구글 브레인 팀이 만들었고 2015년 11월 9일 아파치 2.0 오픈소스 라이선스로 공개되었다[7]. 일종의 다차원 배열인 텐서 내에서의 변수들의 변화흐름을 나타낸다. 이것의 특징은 다음과 같다. 데이터 플로우 그래프를 통한 다양한 표현 방법이 가능하다. 코드 수정 없이 CPU/GPU 모드로 동작할 수 있다. 계산 구조와 목표 함수만 정의하면 자동으로 미분 계산을 처리할 수 있다. Python/C++를 지원하며, SWIG를 통해 다양한 언어 지원이 가능하다. 하지만 Unix계열 OS(Linux/Mac OSX)만 지원하고 GPU 버전은 Linux만 지원하는 것이 단점이다. 윈도우 사용자들은 가상 머신을 이용하거나 도커 툴박스를 설치하여 사용할 수 있다.

III. 태양광발전 시스템 고장 예측

1. Logistic Classification

1	계월	시간	현재발전량	수평발전량	경사발전량	모듈온도	인버터온도	...	레이블
91451	0	15	15	8	11	6	4		0
91452	0	15	15	7	8	5	3		0
91453	0	15	15	7	8	5	3		0
91454	0	15	14	7	5	5	4		0
91455	0	15	14	7	5	5	4		0
91456	0	16	13	5	5	3	3		0
91457	0	16	13	5	5	3	3		0
91458	0	16	12	8	8	4	4		0
91459	0	16	12	8	8	4	4		0
91460	0	16	11	6	12	4	4		0
91471	0	16	11	6	12	4	4		0
91472	0	16	11	6	10	3	4		0
91473	0	16	11	6	10	3	4		0
91474	0	16	10	6	3	3	4		0
91475	0	16	10	6	3	3	4		0
91476	0	16	9	4	5	4	3		0
91477	0	16	9	4	5	4	3		0
91478	0	17	8	4	3	5	3		0
91479	0	17	8	4	3	5	3		0
91480	0	17	7	2	3	4	3		0
91481	0	17	7	2	3	4	3		0
91482	0	17	5	3	3	3	3		0
91483	0	17	5	3	3	3	3		0
91484	0	17	3	3	7	3	3		0
91485	0	17	3	3	7	3	3		0
91486	0	17	2	2	0	4	3		0
91487	0	17	2	2	0	4	3		0

그림 2 태양광발전 모니터링 데이터

그림 2는 약 9만개의 태양광발전 모니터링의 데이터의 cvs파일이다. 에러코드는 에러가 있을 경우 1이고 없을 경우 0으로 나타낸다. 7가지의 조건으로 판단할 때 에러가 없는 경우가 대다수이고 소수의 경우가 에러이다. 그래서 태양광발전 시스템 고장 예측에 사용한 머신러닝 기법은 Logistic Classification이다. 그림 3은 Logistic Classification에서 사용하는 Sigmoid함수의 그래프이다. Logistic은 그림 3과 같이 x값이 증가할

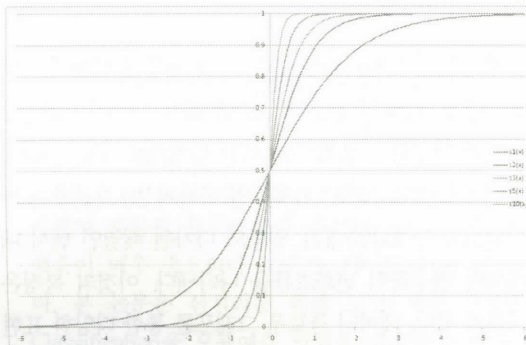


그림 3 Sigmoid 함수의 그래프

수록 y가 1에 수렴한다. 이를 태양광발전 데이터에 접목한다면 소수의 에러 데이터(1)와 다수의 정상 데이터(0)을 Sigmoid로 regression할 수 있다. Logistic Classification은 Sigmoid모양의 Hypothesis를 H(X)라고 선언한다. 그리고 비용을 최소화하기 위하여 아래의 수식을 사용한다.

$$cost(W) = \frac{1}{m} \sum c(H(x), y)$$

$$C(H(x), y) = -y \log(H(x)) - (1-y) \log(1-H(x))$$

cost(W)는 Hypothesis와 data의 비용의 모든 합을 나타낸다. 이 값이 최소값이 될 때 Hypothesis가 정해지는 것으로 판단한다. GD알고리즘을 통하여 최소값을 찾아야 하지만 y가 0일 때와 1일 때 발산하기 때문에 C(H(x), y)의 수식을 이용한다.

2. 텐서플로우를 이용한 고장 예측 결과

Windows 10 환경에서 Anaconda를 이용하여 텐서플로우를 사용하였다. 그림 5는 텐서플로우를 통한 머신러닝 결과화면이다. 러닝 횟수는 총 10만 번이다. 10만 번 동안의 비용, Hypothesis, Classification결과, 정확도가 나타난다. 에러 코드를 제외한 나머지 7가지의 값들을 이용하여 예측

```
100000 0.00387555
Hypothesis:
[[ 1.02786534e-02]
 [ 1.02786534e-02]
 [ 3.95393636e-06]
 ....
 [ 9.25221830e-05]
 [ 2.96149319e-05]
 [ 2.96149319e-05]]
Correct (Y):
[[ 0.]
 [ 0.]
 [ 0.]
 ....
 [ 0.]
 [ 0.]
 [ 0.]]
Accuracy: 0.999956
```

그림 4 텐서플로우의 머신러닝 결과

한 에러 코드의 정확도가 0.999956으로 약 99.9% 일치하였다. 그림 5는 hypothesis와 비용의 러닝 횟수별 변화를 그래프로 보여준다. hypothesis를 나타내는 그래프는 x(나머지값)에 대한 y(에러코드)를 나타내는 그래프이다. 러닝을 거듭할수록 sigmoid의 기울기가 증가하는 것을 확인할 수 있

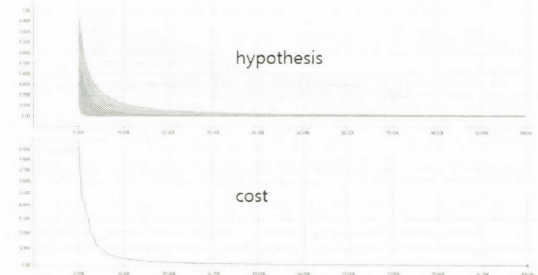


그림 5 텐서플로우의 머신러닝 결과 그래프

다. cost를 나타내는 그래프 역시 10,000번을 기점으로 일정 수렴을 볼 수 있다.

IV. 결론 및 향후 연구

본 논문에서는 Logistic Classification을 이용하여 데이터가 들어왔을 경우 고장 데이터인지 아닌지를 판단하는 방법을 제안한다. 기존 데이터가 한정되어 고장 데이터와 정상 데이터 비율이 비슷하지 않으므로 sigmoid를 이용하여 데이터를 분류하였다. 이를 통해 태양광 발전소의 발전효율을 높일 것이라 기대한다.

현재는 데이터가 약 1년 치뿐이다. 정상인 데이터를 통하여 정상인 데이터 범위가 아닐 경우 고장 데이터를 예측할 수 있다. 하지만 고장 데이터

를 학습하여 판단하는게 아니므로 더 많은 고장 데이터를 수집하여 제대로된 머신러닝을 할 계획이다.

ACKNOWLEDGMENT

이 논문은 2015년 교육부와 한국연구재단의 지역혁신창의인력양성사업의 지원을 받아 수행된 연구임(NRF-2015H1C1A1035548)

참고문헌

- [1] 미국 태양에너지 산업협회 (U.S. Solar Energy Industries Association).
- [2] Hyun Seung Son and R. Young Chul Kim, "Modeling a Photovoltaic Monitoring System based on Maintenance Perspective for New & Renewable Energy", IJCC 2016, AACL 07, pp. 144-147, 2016
- [3] 한국수출입은행 해외경제연구소 (The Export-Import Bank of Korea).
- [4] 김호용, "공간통계기법을 이용한 태양광발전 시설 입지 정확성 향상 방안", 한국지리정보학회지, 13권 2호, pp. 146-156, 2010.
- [5] 제현우, 양오, "지그비 통신을 이용한 태양광 인버터의 원격 모니터링 시스템", 한국정보기술학회논문지, 제10권 제2호, pp. 94-101, 2012
- [6] 제현우, 양오, "스마트폰을 이용한 태양광 인버터의 원격 모니터링 시스템", 한국정보기술학회논문지, 제10권 제1호, pp. 241-249, 2012
- [7] TensorFlow, "<https://www.tensorflow.org/>"