

제지 생산관리시스템의 데이터 자동 수집을 위한 랜덤 포레스트 기반 이상 프로세스 및 원인 분류

손형준¹, 최진영², 김선혁³, 도윤미², 신한솔^{2*}

¹경기대학교, ²한국전자통신연구원, ³국립공주대학교

¹shjmj187@kyonggi.ac.kr, ²{choij0, ydoh, *hansol.shin}@etri.re.kr, ³seonh@kongju.ac.kr

Abnormal Process and Cause Detection using Random Forest Classification for Automated Fault Data Collection in Paper-making MES

Son Hyeong Jun¹, Choi Jin young², Kim Seon Hyeog³, Doh Yoon Mee², Shin Han Sol^{2*}

¹Kyonggi Univ., ²*ETRI, ³Kongju National Univ.

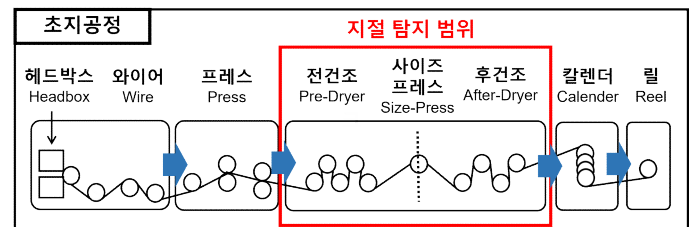
요약

본 연구에서 랜덤 포레스트 기법을 활용해 제지 공정의 지절 프로세스 및 지절 원인을 자동으로 분류하는 모델을 제안한다. 실험 결과, 제안된 모델은 지절 프로세스와 원인 분류에 대한 평균 F1 score가 각각 0.986, 0.975로 나타났다. 생산관리시스템의 지절 데이터 수집을 자동화하여 더 정확하고 효율적인 수집이 가능하고, 구체적인 지절 정보를 제공하여 추후 예방 연구에 기여 가능하다.

I. 서론

공정에 이상이 발생하면 생산 계획에 차질이 생겨, 불필요한 에너지 소비로 생산성이 저해된다[1]. 공정 이상의 원인을 파악하고, 이를 분석하여 예방한다면 이상 발생을 최소화할 수 있다. 이를 위해 첫째로, 이상 데이터가 구체적으로 수집되어야 한다. 전체 공정에서의 이상 발생 여부뿐만 아니라 어떤 공정에서 이상이 발생하는지, 어떤 원인으로 발생하는지에 대한 정보가 필요하다. 둘째로, 정확하고 효율적으로 수집되어야 한다. 인력에 의존하지 않는 자동화된 분류 기술이 필요하다. 본 연구에서는 제지 업종의 주요 공정 이상인 '지절'을 중심으로 랜덤 포레스트(Random Forest) 기법을 활용해 생산관리시스템의 지절 데이터 자동 수집을 구현한다.

지공정의 전체적인 프로세스이다. 지절 발생 사례 수가 적은 프로세스들은 분석 대상에서 제외하고, 전체 지절 발생의 93.4%를 차지하는 전건조, 사이즈프레스, 후건조 프로세스에 대해 지절 탐지를 진행했다.



〈그림 1〉 초지공정과 지절 탐지 범위

II. 본론

2.1 지절 데이터 자동 분류 및 수집 시스템의 필요성

지절 데이터는 생산관리시스템(Manufacturing Execution System, MES)을 통해 관리되나, 지절이 발생한 프로세스와 원인은 작업자의 경험과 판단을 기반으로 분류된다. 즉, 작업자 주관에 따라 분류가 달라질 수 있어 일관성이 떨어지고, 오류 발생 가능성이 크다. 또한 시간과 비용이 많이 소모되고, 누락과 입력 지연이 발생할 수 있으므로 비효율적이다.

기계학습 모델을 이용하여 전문가 개입 없이 자동으로 공정 데이터를 분류하면 시간과 비용의 절감 뿐만 아니라, 데이터의 일관성과 정확성이 확보되어 표준화가 가능하다. 그리고 분류 모델의 고도화를 통해, 실시간 지절 모니터링 및 지절 발생 시점 예측과 같은 시스템 확장이 가능하다.

2.2 지절 탐지 범위 및 프로세스별 원인 비율

제지 공정은 조성공정, 초지공정, 코팅공정, 완정공정으로 이루어진다. 이 중 초지공정은 원료들을 탈수, 건조하여 종이의 형태를 만드는 공정이다. 이 과정에서 스팀 및 열에너지를 많이 소비한다. 따라서 초지공정에서 발생하는 지절을 개선한다면 많은 에너지를 절감할 수 있다. 그림 1은 초

지절이란 종이가 생산, 가공, 인쇄 등의 공정 중에 찢어지는 현상을 말한다[2]. 실험에 사용한 공정 데이터를 기준으로 지절은 전체 공정 시간의 0.33%를 차지했다. 또한 한 달에 평균 5.8번 발생하고, 한 번 발생할 때마다 평균 25.7분 동안 무효 가동을 일으켰다. 지절 원인은 크게 다섯 가지로 나눌 수 있다. 종이의 꼬트머리가 찢어지는 '귀상', 종이에 작은 구멍이 생기는 '홀', 설비에 이물질이 낙하하거나 혼입되는 '이물질 혼입', 고무 블레이드 같은 부품에 문제가 발생했을 때 의도적으로 진행되는 '부품 교체', 센서의 오작동, 기계 오입력 등 의도치 않은 기계 문제로 인한 '기계 오류'가 있다. 프로세스에 따른 원인별 발생 비율은 표 1과 같다.

〈표 1〉 지절 프로세스별 지절 원인 비율 [단위: %]

프로세스	지절 원인					합계
	귀상	Hole	이물질 혼입	부품 교체	기계 오류	
SizePress	55.7	11.0	11.7	13.6	7.9	33.7
PreDryer	21.1	51.8	15.0	0	12.1	53.2
AfterDryer	82.7	17.3	0	0	0	13.1
합계	40.9	33.5	11.9	4.6	9.1	100.0

2.3 랜덤 포레스트를 활용한 분류 모델

랜덤 포레스트는 여러 개의 의사결정나무를 배깅 방식으로 결합하여 만든 분류 알고리즘이다. 개별 트리가 독립적으로 학습한 결과를 종합함으로써, 복잡한 제지 공정에서 발생하는 지절 원인을 효과적으로 분류할 수 있다. 또한 과적합을 방지하고 변수의 중요도를 평가하는 데 유용하다[3]. 본 연구에서는 랜덤 포레스트 모델을 이용하여 피쳐 중요도가 높은 공정 태그를 독립변수로 활용해 지절 프로세스와 지절 원인을 분류하는 알고리즘을 제안한다. 이를 통해, 어떤 프로세스에서 지절이 발생했는지를 탐지하고, 더 나아가 해당 프로세스의 지절 원인을 파악하고자 한다.

2022년 1월 1일부터 11월 30일까지 1분 단위로 측정되어, 79,386개(분)의 정상 데이터와 1,574개(분)의 지절 데이터로 이루어진 MES 데이터를 사용했다. 언더 샘플링을 진행하여 정상-지절 데이터 간의 불균형을 처리하고 실험을 진행했다. 실험 1의 종속변수는 정상상태(OK)와 지절이 발생하는 각 프로세스(SizePress, PreDryer, AfterDryer)이며, 랜덤 포레스트 모델의 중요도 상위 10개의 피쳐를 독립변수로 선정했다. 실험 2는 실험 1의 독립변수를 사용하여 정상상태(OK)와 15종류의 프로세스별 지절 원인까지 총 16종류의 종속변수를 분류했다. 분류 모델의 독립변수에 대한 설명은 표 2와 같다.

〈표 2〉 분류 모델에서 사용한 독립변수

피쳐명	설명	피쳐 중요도(%)	
		실험 1	실험 2
X1	작업 횟수	19.96 (1위)	15.31 (1위)
X2	가동 시간 (분)	19.01 (2위)	13.6 (3위)
X3	릴(Reel)의 길이	9.02 (3위)	13.92 (2위)
X4	생산 속도	8.71 (4위)	6.82 (10위)
X5	평량별 작업 횟수	8.66 (5위)	8.88 (5위)
X6	종이의 밝기	8.22 (6위)	10.11 (4위)
X7	종이의 두께	7.78 (7위)	8.72 (6위)
X8	종이의 백색도	6.98 (8위)	7.78 (8위)
X9	종이의 평량 (무게)	6.06 (9위)	7.10 (9위)
X10	회색수의 양	5.59 (10위)	7.84 (7위)

지절 프로세스와 프로세스별 지절 원인을 분류한 혼동 행렬은 그림 2 및 그림 3과 같으며, 분류 모델별 정밀도와 재현율이 유사하다. 각 모델별 F1 Score는 표 3 및 표 4와 같다. 예측 결과, 지절 프로세스 분류와 지절 원인 분류에 대한 평균 F1 score는 0.986 및 0.975로 계산되었다.

실제	98.5% (473)	0.6% (3)	0.8% (4)	0.0% (0)	98.5%
	1.3% (3)	98.7% (225)	0.0% (0)	0.0% (0)	98.7%
	1.8% (6)	0.3% (1)	97.9% (326)	0.0% (0)	97.9%
	1.2% (1)	0.0% (0)	0.0% (0)	98.8% (82)	98.8%
	97.9%	98.3%	98.8%	100.0%	
예측					

〈그림 2〉 지절 프로세스 분류 혼동행렬

〈그림 3〉 지절 원인 분류 혼동행렬 [단위: %]

〈표 3〉 지절 프로세스 분류 F1 Score

종속변수	OK	SizePress	PreDryer	AfterDryer
F1 score	0.982	0.985	0.983	0.994

〈표 4〉 지절 원인 분류 F1 Score

프로세스	OK	귀상	Hole	이물질 혼입	부품 교체	기계 오류
SizePress	0.982	0.972	0.966	0.974	0.981	0.968
PreDryer		0.980	0.974	0.992	(사레없음)	0.977
AfterDryer		0.978	0.960	(사레없음)	(사레없음)	(사레없음)

III. 결론

본 연구에서는 제지 공정의 주요 이상 현상인 지절을 중심으로, 랜덤 포레스트 기법을 활용한 지절 데이터 자동 분류 모델을 제안했다. 실험 결과, 제안된 분류 모델은 지절이 발생한 프로세스와 원인별로 각각 평균 0.986 및 0.975로 높은 F1 score를 기록했다. 이는 생산관리시스템에서 수기로 진행되었던 지절 데이터 수집을 자동화하여 지절 데이터의 일관성 및 정확성을 확보할 수 있음을 시사한다. 또한 구체적인 지절 정보를 제공하여 지절에 효율적으로 대처하고 예방할 수 있다. 추후 본 연구를 기반으로, 실시간 지절 모니터링 및 지절 발생 시점 예측을 연구할 계획이다.

ACKNOWLEDGMENT

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구 과제입니다. (No. 20202020800290)

참 고 문 헌

[1] 김선혁, 도윤미, 신영미, 권순현, 이좌형, 이상금, 허태욱. (2022). 공장 에너지관리시스템(FEMS)의 공정 이상 모니터링을 위한 신경망 기반 이벤트 분류 알고리즘. *한국통신학회 학술대회논문집*, 제주.

[2] 박중문. (2021). 지절에 관련된 파괴인성의 이해, 측정과 활용. *펄프·종이기술*, 53(1), 5-14.

[3] Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32.