

**잠재 소비자 예측 및 관심사 클러스터링 기반
생성형 AI 프롬프트 활용 임산물별 맞춤 활성화 전략**

기획 개요	본 프로젝트는 임산물의 종류별로 잠재 고객을 파악하여, 잠재 고객의 특성과 관심사를 고려한 임산물별 맞춤형 활성화 방안을 도출하는 것을 목표로 한다. 이를 위해, 1단계에서는 임산물별 사람들의 관심 정도와 특성을 분석하고, 2단계에서는 임산물 종류별 잠재 고객을 예측하기 위한 분류 예측 모델을 개발하여 실제 잠재 고객을 분류한다. 이어, 3단계에서는 잠재 고객의 관심사를 검색/블로그/뉴스 기사 등의 글을 기반으로 클러스터링하여 추출하고, 4단계에서는 잠재 고객들의 관심사, 연령대, 직업 등을 고려하여 임산물 종류별 인사이트를 추출한다. 이때 추출한 인사이트는 생성형 AI의 프롬프트로 입력되며, 이를 기반으로 구체적이고 실현 가능하면서도 창의적인 임산물별 활성화 방안을 제시한다. 이러한 프로세스는 기술적, 산업적, 정책적 측면에서 중요한 의미를 가지며, 임산물 산업의 활성화를 촉진하는 동시에 수요와 공급의 불균형 문제를 완화할 수 있는 계기가 될 것으로 기대한다.		
활용데이터 정보	출처	제공기관명	데이터명
	1. https://kfss.forest.go.kr/stat/	산림통계시스템	임산물 생산비조사
	2. www.kfss.forest.go.kr	산림청 산림통계시스템	2023년 임업경영실태조사 주요결과 인포그래픽
	3. https://www.kofpi.or.kr/index.do	한국임업진흥원	임가소득의 구성요소별 증감률
	4. www.bigdata-forest.kr	산림빅데이터 플랫폼	2019~2020년 성/연령/지역 등 이용자 분류별 6대 임산물 관심도 정보
	5. www.bigdata-forest.kr	산림빅데이터 플랫폼	농산물 소비인구 정보
	6. www.bigdata-forest.kr	산림빅데이터 플랫폼	임산물 관심인구 정보

1) 추진배경 및 필요성

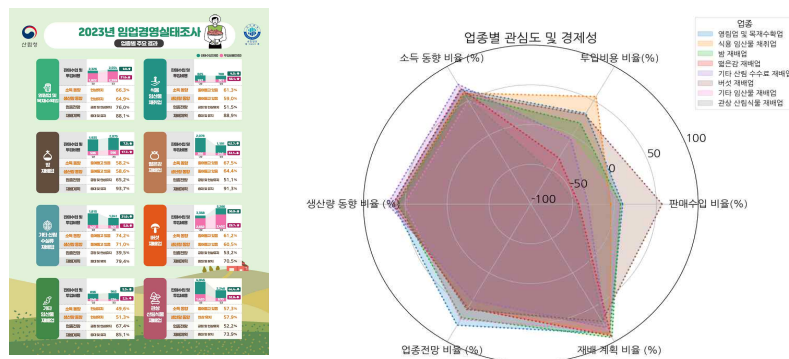


그림 1 업종별 관심도 및 경제성



그림 2 임가소득의 구성요소별 증감률

임산물	밤	대추	호두	뽕은감	오미자	구기자	산수유	복분자	산딸기	오갈피	헛개나무
순수익	-1,541,235	-5,559,958	-1,366,896	3,449,948	-13,546,515	-24,213,413	-9,786,895	-13,864,315	-35,237,305	-9,454,434	-7,796,054

표 1 2023년 임산물별 1ha당 소득분석(단위: 평가액(원))

산림청 산림통계시스템의 '2023년 임업경영실태조사' 인포그래픽 및 이를 레이더 차트화한 그림 1에 따르면, 대부분의 업종에서 임업소득 및 생산량이 줄어 들고 있음을 알 수 있다. 더불어 한국임업진흥원의 '임가소득의 구성요소별 증감률' 그래프인 그림 2에서 알 수 있듯이 2019년에 비해 2023년 임업소득은 약 2,913천원 줄어든 상태로 하락세를 보인다. 또한 표 1과 같이 산림통계시스템의 '임산물 생산비조사' 데이터를 보면, 뽕은감을 제외한 임산물별 순수익이 모두 음수, 즉 적자임을 알 수 있다. 임산물의 소비를 촉진하는 사업[1,2]도 진행되고 있으나 이러한 문제점이 지속되는 이유 중 하나는 임산물에 대한 수요가 적어지기 때문이다[3]. 이는 결국 임산물에 대한 수요와 공급의 불균형을 초래하며, 이는 생산자뿐만 아니라, 생산량이 적어지게 되면 소비자 또한 가격 측면 등에서 피해를 볼 수 밖에 없는 악순환 구조[4]를 가지게 된다.

또한, 현재 임산물별로 뚜렷하게 다른 특성을 보임을 알 수 있다. 그림 1을 보면, 밤 재배업은 판매 수입 비율이 7.2% 증가하고 재배 계획 비율이 93.7%로 나타나, 소비자들에게 지속적으로 선호받고 있으며 안정적인 수요를 유지하고 있음을 보여주었다. 반면, 뽕은감 재배업은 판매수입 비율이 -42.7%로 크게 감소했지만 재배 계획 비율이 91.3%로 유지되고 있어, 기존 소비자층은 안정적이거나 새로운 소비자 유입이 어려운 구조임을 확인할 수 있었다. 기타 산림 수수로 재배업은 재배 계획 비율이 79.4%로 낮은 수준을 보이며 소비자 관심이 줄고 있음을 시사하였다. 이와 같은 결과는 임산물의 경제적 성과와 소비자 선호도가 품목별로 차이가 크다는 점을 잘 설명한다.



그림 3 프로젝트 프레임워크

따라서 임산물에 대한 수요, 공급을 맞추기 위해서는 임산물 종류별로 상이한 특성을 고려해 활성화 방안을 세울 필요가 있다. 이를 위하여 본 프로젝트에서는 그림 3과 같은 분석 프로세스를 통해 활성화 방안 도출 방법을 제안하고자 한다. 구체적으로, 각 임산물의 종류에 따른 잠재 고객을 예측하기 위한 모델을 개발하며, 예측된 잠재 고객의 관심사와 연령 등의 특성을 클러스터링과 같은 기법을 통해 분석하여 인사이트를 도출한다. 이 인사이트를 활용하여 프롬프트를 제시하고, 제시한 프롬프트를

생성형 AI에 입력하여 임산물의 종류별로 구체적이면서도 현실적인 활성화 방안을 발굴한다.

이러한 접근법은 소비자의 관심을 유도하고, 임산물의 소비를 활성화하여 수요와 공급의 불균형 문제를 해결하는 중요한 돌파구로 작용할 수 있다. 이때 맞춤형 프롬프트 설계의 효과를 입증하는 논문 [5]과 같이 생성형 AI의 활용은 기존 대비 임산물별 특성에 따라 최적화된 활성화 방안을 도출할 수 있다. 이는 각 임산물의 시장 가치를 극대화하는 데에도 기여할 것이다.

2) 아이디어 기획 세부 내용

본 프로젝트의 주요 과제는 임산물에 대한 사람들의 관심사를 분석하고, 이를 토대로 임산물 구매 가능성 인구를 예측해 그들의 관심사를 파악하고 맞춤형 활성화 방안을 제안하는 것이다. 이를 위하여 아래와 같은 단계로 분석을 진행한다.

1단계: 임산물별 사람들의 관심 정도 및 특성 분석

1단계는 임산물의 종류에 따른 사람들의 관심 정도와 관심 있는 사람들의 특징을 분석해 인사이트를 얻는 과정이다.

‘2023년 임업경영실태조사’ 자료를 바탕으로 임산물 종류별 관심도를 분석한 결과, 식용 임산물 채취업이 두 번째로 높은 관심도를 기록한 것으로 나타났다. 이는 관심도를 계산하기 위해 아래와 같은 식을 사용한 결과이다.

$$\text{관심도 점수} = \left(\frac{\text{판매수입 (백만원)}}{\text{투입비용 (백만원)}} \times 0.25 \right) + (\text{소득동향비율 (\%)} \times 0.25) + (\text{생산량동향비율 (\%)} \times 0.25) + (\text{업종전망비율 (\%)} \times 0.25)$$

이 식은 임산물 업종별로 경제적 성과와 시장에서의 위치를 종합적으로 평가하기 위해 사용되었다. 판매수입 대비 투입비용 비율은 해당 업종의 수익성을 나타내며, 소득 동향, 생산량 동향, 업종 전망 비율은 시장에서의 안정성과 성장 가능성을 평가하는 주요 요소이며, 각 항목에 동일한 가중치(0.25)를 부여하였다.

분석 결과, 그림 4를 보면 식용 임산물 채취업은 안정적인 투입비용 대비 판매수입 비율과 균형 잡힌 소득 및 생산량 동향 비율로 인해 두 번째로 높은 관심도를 기록한 것을 확인할 수 있다. 이는 해당 업종이 안정적인 경제적 성과를 바탕으로 시장 내에서 중요한 위치를 차지하고 있음을 나타낸다. 이러한 품목은 고부가가치 상품 개발 및 시장 확대 전략을 통해 추가적인 성장을 도모할 여지가 있다. 밤 재배업과 뽕은감 재배업은 중간 정도의 관심도를 기록한 품목들로 나타났다. 이들 품목은 작년대비 관심도 점수가 소폭 상승하거나 유지된 것으로 보인다. 이는 소득 및 생산량 동향에서 부분적으로 긍정적인 흐름을 보였기 때문으로 해석된다. 반면, 기타 산림 수수료 재배업과 관상 산림수목 재배업은 가장 낮은 관심도를 기록하였다. 이들 품목은 소득 및 생산량 동향 비율이 감소세를 보였으며, 업종 전망 비율도 상대적으로 낮았다. 이는 시장에서 경제적 성과가 떨어지고 경쟁력이 약화된 상태를 반영한다.

결론적으로, 임산물 종류별 관심도 분석은 각 품목의 경제적 성과와 시장 위치의 차이를 반영하고 있다. 관심도가 높은 품목은 안정적인 성과를 유지하며 추가적인 성장을 도모해야 하며, 관심도가 낮은 품목은 구조적 문제를 개선하고 경쟁력을 강화하기 위한 노력이 필요하다. 각 항목의 중요도를 균등하게 고려한 관심도 점수 계산 방식을 통해, 품목별 맞춤형 전략 수립의 필요성을 확인할 수 있다.

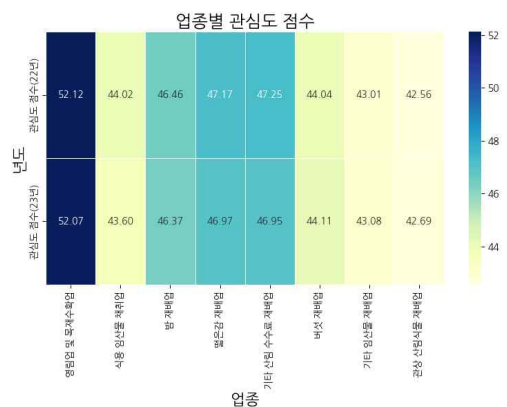


그림 4 업종별 관심도 점수

그림 5는 '임산물 관심인구 정보' 데이터에서 임산물구분명과 월별 정보획득 데이터를 활용하여 구매

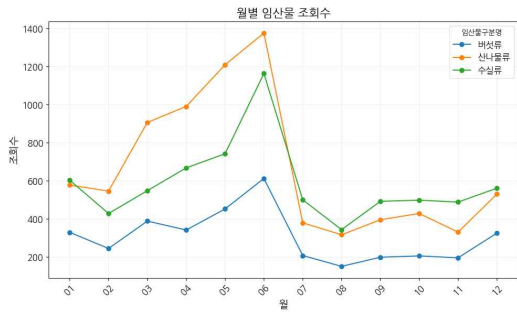


그림 5 월별 임산물 조회수

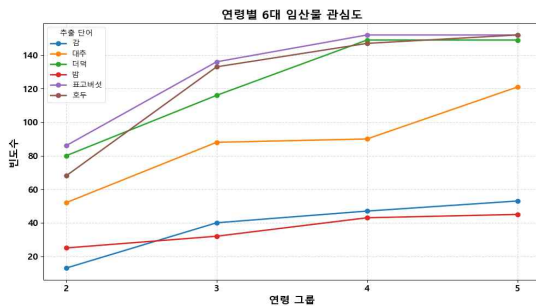


그림 6 연령별 임산물 관심도

자의 계절적 관심도를 파악한 그래프이다. 구매자의 관심도는 임산물의 종류와 월별 시즌에 따라 차이를 보였다. 이는 구매자의 계절적 소비 패턴이 임산물 종류별로 다르게 나타남을 의미하며, 임산물별 맞춤형 활성화 방안이 필요하는 것을 알 수 있다.

그림 6는 산림 빅데이터 거래소의 '2019년~2020년 성/연령/지역 등 이용자 분류별 6대 임산물 관심도 정보' 데이터를 활용하여, 연령 그룹별 표고버섯, 호두, 감, 밤, 대추 등에 대한 관심도 변화를 분석한 그래프이다. 연령 그룹에 따라 임산물 종류별로 관심도가 상이하게 나타나는 것은 품목마다 고유한 경제적 성과와 시장에서의 위치가 다르다는 것을 의미하며, 품목 특성에 맞춘 차별화된 전략이 필요하는 것을 시사한다.

2단계: 임산물 구매가능성 인구 예측

2단계에서는 1단계의 관심인구 특성을 기반으로 임산물 구매 가능성이 있는 인구를 예측한다. 이를 위해 산림 빅데이터 거래소의 '임산물 관심인구 정보' 일부와 '농산물 소비인구 정보' 데이터를 사용한다.

우선, 두 데이터를 합쳐 분류 예측 모델을 학습 및 평가하기 위한 데이터셋을 만들어준다. '농산물 소비인구 정보'에는 임산물에 속하지 않는 농산물 데이터도 포함되어 있기 때문에 해당 데이터를 이용하면 임산물을 구매할 인구를 예(1), 아니오(0)로 분류하도록 학습시킬 수 있다. 여기에 임산물의 종류와 데이터 증강을 위해 더 많은 임산물 정보가 담긴 '임산물 관심인구 정보' 데이터 행의 50%를 랜덤 언더샘플링한 후 농산물 데이터와 결합해 사용한다. 이를 통해, 데이터 확보와 데이터 불균형 문제 해결을 동시에 달성한다. 참고로, 나머지 절반의 행에 대해서는 다음 단계인 3단계에서 구매 가능성 인구를 예측하는 데 사용한다.

'농산물 소비인구 정보' 데이터의 경우 임산물명이 따로 기재되지 않았기 때문에, 제품명 컬럼에서 임산물명을 추출한다. 예를 들어 '감자탕'의 경우 '감' 글자가 들어갔다고 임산물명을 '감'으로 잘못 인식되지 않게 예외처리를 한다.

임산물 종류	해당 임산물명
견과류	호두, 밤, 도토리, 잣, 개암, 은행
버섯류	표고버섯, 꽃송이버섯, 송이버섯, 석이버섯, 싸리버섯, 능이버섯, 목이버섯, 복령버섯
나물류	산마늘, 고사리, 눈개승마, 곤드레(고려엉귀), 참나물, 어수리, 두릅, 취나물, 고비, 원추리, 곰취, 들메나무순
열매/과일류	대추, 머루, 복분자, 산딸기, 돌배, 감, 석류, 다래
뿌리/줄기류	죽순, 도라지, 더덕

표 2 임산물 종류 분류

이후, 변수 간 상관분석으로 다중공선성을 확인하고 최종 독립변수를 선별한다. 종속변수에 해당하는 임산물은 각 특성에 맞게 모델을 일반화하고 그 성능을 향상시키기 위해서 표 2와 같이 그 종류로 구분한다.

각 임산물 종류별 구매 여부 파악을 위해분류 예측 모델링을 진행한다. 이때 사용한 모델은 여러 의사결정트리를 앙상블하여 안정적인 성능을 내는 RandomForest[6], 그래디언트 부스팅 기반으로 속도

와 예측력을 확보한 XGBoost[7], 그래디언트 부스팅을 정규화해 예측 불확실성을 줄인 NGBoost[8], 비선형 패턴과 복잡한 데이터를 효과적으로 학습하는 3-layer MLP[9] 모델이다.

각 모델의 성능을 비교 및 검증하기 위해 성능 지표로 Accuracy, MSE, RAE를 사용한다. 정확도는 예측이 얼마나 맞았는지를 보여주는 지표로, 모델의 전체 예측값 중 정답을 맞춘 비율을 나타낸다. MSE는 예측값과 실제 값 사이의 차이를 제곱하여 평균을 낸 값이며, RAE는 예측 오차를 상대적으로 평가한 지표이다. 이 둘은 값이 작을수록 오차가 적으므로 모델의 예측이 실제값과 가까움을 의미한다.

이를 통해 모델들의 성능을 비교하여 각 임산물 종류별 가장 성능이 좋은 최적의 모델을 선택한다. 임산물 종류에 따라 별개로 구축된 모델은 각각 다른 학습 가중치를 가진다는 의미이며, 임산물 데이터 전체를 한 번에 학습시킨 통합 모델에 비해 임산물 종류 특성을 기반으로 학습한 정밀한 맞춤형 예측 가능하다는 장점을 가진다.

임산물 종류	모델	정확도	MSE	RAE
견과류	RandomForest	0.785507	0.214493	0.429595
	XGBoost	0.805797	0.194203	0.388755
	NGBoost	0.782609	0.217391	0.435173
	3-layer MLP	0.696618	0.303382	0.607308
버섯류	RandomForest	0.779141	0.220859	0.442673
	XGBoost	0.805797	0.194203	0.388755
	NGBoost	0.774759	0.225241	0.450541
	3-layer MLP	0.718668	0.281332	0.562737
나물류	RandomForest	0.927905	0.072095	0.144266
	XGBoost	0.974555	0.025445	0.051002
	NGBoost	0.914334	0.085666	0.171705
	3-layer MLP	0.875318	0.124682	0.249908
열매/과일류	RandomForest	0.657032	0.342968	0.686102
	XGBoost	0.672131	0.327869	0.656238
	NGBoost	0.661346	0.338654	0.677825
	3-layer MLP	0.631579	0.368421	0.737404
뿌리/줄기류	RandomForest	0.880295	0.119705	0.239548
	XGBoost	0.885820	0.114180	0.228399
	NGBoost	0.661346	0.338654	0.677825
	3-layer MLP	0.751381	0.248619	0.497320

표 3 임산물 종류별 관심 여부 분류 예측 모델 성능 결과

학습 결과는 표 3과 같다. 모든 임산물 종류에서 XGBoost 모델이 가장 좋은 성능을 보여주어 이 모델을 각 임산물 종류별 최종 모델로 결정하여 3단계의 임산물 구매 가능성 여부를 예측할 때 사용하도록 한다.

3단계: 임산물에 대한 관심사 토픽 모델링

3단계에서는 산림 빅데이터 거래소의 '임산물 관심인구 정보' 데이터를 사용해 임산물에 대한 관심사를 도출한다. 검색해 들어간 URL의 제목을 기준으로 토픽 모델링을 통해 이들의 관심사를 클러스터링

	Representation	topic
0	[만들기, 무침, 맛있게, 요리, 묵무침, 볶음, 졸은, 레시피, 나물, 만들기, ...]	요리/레시피
1	[나무, 씨앗, 재배, 자연산, 열매, 과일, 산행, 채취, 꽃이, 홍천, 집에서, ...]	취미
2	[효능, 좋은, 면역력, 음식, 먹는법, 배즙, 건강, 효능과, 높이는, 당뇨에, ...]	건강
3	[판매, 가격, 파는곳, 산업, 나물, 지원, 수출, 경북, 2020년, 수출길, ...]	판매정보
4	[2tv생생정보통맛집오늘, tv, 소마면, 갈비조림, 바지락술밥, 정식, 통갈치조림...]	TV매체

그림 7 BERTopic으로 토픽 모델링한 결과

끼리 묶어 총 5가지 주제의 관심사로 분류하였다. 요리/레시피, 취미, 건강, 판매정보, TV매체가 이에 해당한다. 3단계의 결과는 4단계에서 활성화 방안을 세울 때 고려할 중요한 정보가 된다.

4단계: 생성형 AI를 활용한 관심사 클러스터별 임산물 활성화 방안 모색

4단계는 2단계의 분류 예측 모델을 이용해 임산물 잠재적 구매 가능성이 높은 인구를 추출하고, 3단계에서 도출한 임산물 관련 관심사를 포함한 연령, 지역 등 여러 특성을 고려해 임산물 종류별 활성화 방안을 모색하는 단계이다.

우선 2단계에서 생성한 XGBoost 기반의 분류 예측 모델을 토대로 사람들 중 임산물에 대한 잠재적 구매 가능성이 높은 인구와 그렇지 않은 인구를 분류하고 둘 간의 특성을 비교한다. 이때 사용하는 데이터는 2단계에서 사용하지 않은 나머지 '임산물 관심인구 정보' 데이터이다. 이는 검색/블로그/뉴스기사 등의 사이트에서 임산물에 관심을 보인 사람들의 정보를 담고 있으며, 이들 중 잠재적 구매 가능성이 높은 인구를 추출한다. 총 9,333개의 데이터 중 6,253개가 예측값이 1인 잠재적 임산물 구매 인구로 분류됐다.

이들을 대상으로 데이터 시각화를 통해 그들의 관심 특성 및 임산물별 인사이트를 추출한다. 추출한 대표 인사이트는 다음과 같다.

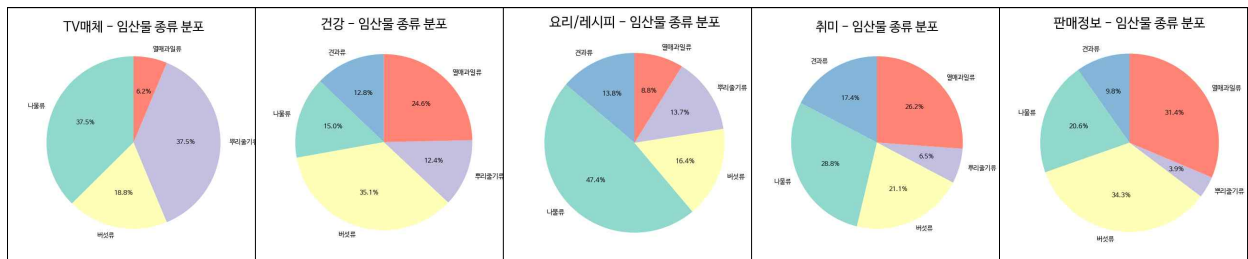


그림 8 관심사별 임산물 종류 분포

[Insight 1] TV매체에 관심 있는 사람들은 나물류와 뿌리/줄기류, 건강에 관심 있는 사람들은 버섯류, 요리/레시피에 관심 있는 사람들은 나물류, 취미에 관심 있는 사람들은 나물류와 열매/과일류, 판매정보에 관심 있는 사람들은 버섯류와 열매/과일류에 관심이 큰 것을 확인할 수 있다. 이를 통해, 나물류의 임산물을 더 흥행시키기 위해서는 다양한 나물 관련 요리/레시피를 알려주거나, 버섯류의 임산물을 흥행시키기 위해서는 건강에 접목시킨 활성화 방안을 모색해보는 것 등을 고려해볼 수 있다.

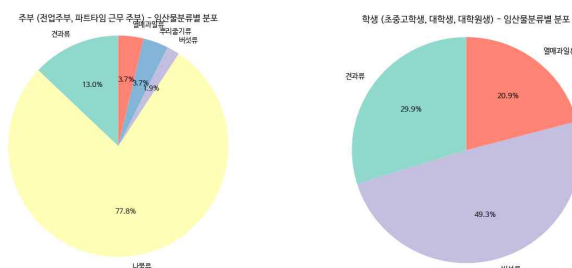


그림 9 직업별 관심 임산물 종류 분포

[Insight 2] 직업별 관심 임산물 종류를 분석한 결과, 주부의 경우 나물류에 압도적으로 많은 관심을

가지고 있고, 학생의 경우 나물보다는 견과류나 버섯류, 열매/과일류에 주로 관심이 있는 것을 확인할 수 있다.

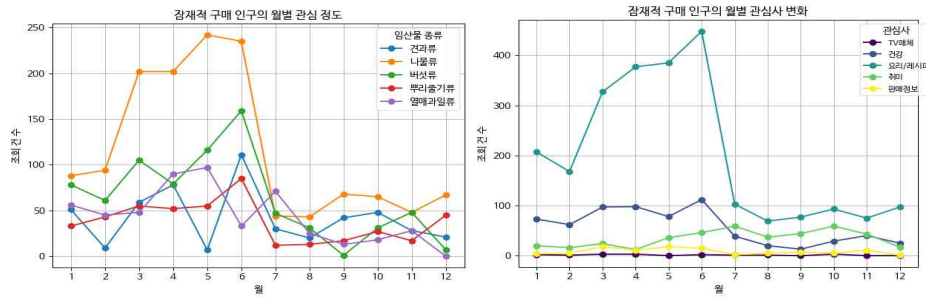


그림 10 잠재적 구매 가능성이 높은 인구의 월별 수요 정도 및 관심사 변화

[Insight 3] 임산물에 대한 월별 관심 정도를 분석한 결과, 그림 10의 좌측 그래프와 같이 봄철과 초여름까지의 기간에 임산물에 대한 관심이 큼을 알 수 있다. 또한 우측 그래프와 같이 관심사별로 그 관심 정도의 월별 추이를 분석한 결과, 요리/레시피에 대한 관심도는 봄과 여름까지가 가장 크며, 가을에 접어들면서 급격히 줄어드는 모양새를 띤다. 건강에 대한 관심도도 비슷한 추이를 보이나, 취미의 경우 여름과 가을 및 초겨울까지 비교적 관심도가 높게 나타남을 알 수 있다.

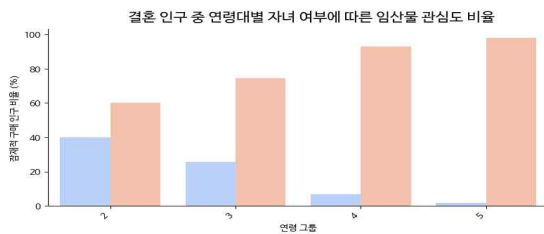


그림 11 결혼 인구 중 연령대별 자녀 여부에 따른 임산물 관심도 비율



그림 12 연령대별 관심사 비율

[Insight 4] 자녀 여부에 따른 임산물 관심도 비율을 조사한 결과(그림 11), 결혼한 사람 중 자녀를 가진 부모의 경우 그렇지 않은 인구보다 임산물에 대한 관심이 훨씬 큰 것으로 나타났다. 특히, 20대 부모보다 30, 40, 50대로 갈수록 그 격차가 더욱 커지는 것을 확인할 수 있다.

[Insight 5] 연령 그룹별 관심사를 조사한 결과(그림 12), 전연령대에서 임산물을 흔하게 접할 수 있는 임산물 요리/레시피에 가장 큰 관심을 가지고 있으며, 20대, 40대의 경우 임산물을 이용한 건강에도 비교적 큰 관심을 가지고 있는 것을 확인할 수 있다.

앞서 얻은 5가지의 인사이트를 기반으로, 생성형 AI 중 가장 성능이 뛰어나며 잘 알려진 ChatGPT[13]를 활용하여 각 임산물 종류별 활성화 방안을 제안한다.

프롬프트
아래 insight들을 고려한 각 임산물 종류에 따른 활성화 방안을 현실적이고 구체적으로 제안해줘.
임산물 종류: 견과류, 나물류, 버섯류, 뿌리/줄기류, 열매/과일류
[Insight 1] TV매체에 관심 있는 사람들은 나물류와 뿌리/줄기류, 건강에 관심 있는 사람들은 버섯류, 요리/레시피에 관심 있는 사람들은 나물류, 취미에 관심 있는 사람들은 나물류와 열매/과일류, 판매정보에 관심 있는 사람들은 버섯류와 열매/과일류에 관심이 큼 [Insight 2] 주부에게는 나물류, 학생들은 견과류나 버섯류, 열매/과일류에 관심이 있음 [Insight 3] 봄철과 초여름까지가 임산물에

대한 관심이 큼 [Insight 4] 결혼한 사람 중, 자녀 있는 사람이 임신물 구매를 훨씬 더 많이 함. 특히, 20대 부모보다 30, 40, 50대로 갈수록 그 격차가 더욱 커짐 [Insight 5] 전연령에서 요리/레시피에 가장 큰 관심. 20대, 40대의 경우 임신물을 이용한 건강에도 관심이 있음

표 4 생성형AI 입력 프롬프트

이러한 프로세스는 ChatGPT에 입력한 프롬프트는 임신물별 정밀한 분석을 통해 뽑아낸 인사이트를 활용하여 작성했다는 점에서 차별성을 갖는다. 또한, 실행 가능한 활성화 방안을 추출하는 것이 중요하다 판단하여 현실적이고 구체적으로 제안하라는 문구를 함께 입력한다. 이를 위해 사용한 프롬프트의 내용은 표 4와 같다.

<p>1. 건과류 활성화 방안</p> <ul style="list-style-type: none"> • 학생층과 간편성 강조 <ul style="list-style-type: none"> • 학생들의 관심을 반영하여, 건과류를 활용한 간편 간식 제품을 개발하고, 대학가와 청소년이 자주 찾는 공간에서 무료 시식 행사를 진행. • SNS 플랫폼을 활용한 "건과류로 만드는 5분 간식" 콘텐츠 제작. • 패키지 디자인에 '공부할 때 최고의 에너지 보충' 메시지 추가. • 건강과 연계된 마케팅 <ul style="list-style-type: none"> • 건과류의 영양적 가치와 집중력 향상 효과를 강조하는 건강 캠페인 진행. • 20~40대가 주로 사용하는 피트니스 앱이나 헬스 관련 매체에 맞춤형 광고 송출. 	<p>2. 나물류 활성화 방안</p> <ul style="list-style-type: none"> • 요리 및 레시피 중심 접근 <ul style="list-style-type: none"> • 전연령층의 요리 관심을 반영하여 나물 요리법 강좌 또는 유튜브 채널 운영. • 간단히 조리 가능한 '손질된 나물 키트'를 판매하고, QR코드로 연결되는 레시피 제공. • TV 매체 활용 <ul style="list-style-type: none"> • TV 요리 프로그램과 협업하여 나물 활용 레시피를 자주 노출. • 주부층을 겨냥해 홈쇼핑에서 '봄철 인기 나물 세트' 기획. • 취미와 연계 <ul style="list-style-type: none"> • "나물로 만드는 색다른 요리"를 주제로 한 쿠킹 클래스 운영. • 자연 체험 프로그램과 연계해 가족 단위의 나물 채집 행사 기획.
<p>3. 버섯류 활성화 방안</p> <ul style="list-style-type: none"> • 건강 이미지 강조 <ul style="list-style-type: none"> • 20대와 40대를 타겟으로 면역력 강화와 건강을 주제로 한 마케팅 진행. • 슈퍼푸드로서의 버섯 효능을 강조하는 캠페인 진행 (예: '면역력을 위한 매일 한 끼 버섯'). • 판매정보 제공 및 소비자 교육 <ul style="list-style-type: none"> • 온라인 판매 플랫폼에 버섯의 효능과 요리법을 상세히 설명. • 판매용 버섯 패키지에 '효능과 활용법 카드'를 포함. • 학생층과 맞물려 '버섯 키우기 키트'를 개발하여 판매. 	<p>4. 뿌리/줄기류 활성화 방안</p> <ul style="list-style-type: none"> • TV 매체와 건강 관심층 타겟팅 <ul style="list-style-type: none"> • 뿌리/줄기류를 활용한 건강주스, 간단 요리법을 TV 매체와 유튜브 채널에서 홍보. • 결혼 후 자녀가 있는 30~50대를 겨냥해 건강 효능을 강조한 '가족 건강 프로젝트' 진행. • 봄철 캠페인 <ul style="list-style-type: none"> • 봄철에 뿌리/줄기류를 활용한 디톡스 프로그램과 연계된 상품 판매. • 농가와 협력해 '봄맞이 뿌리/줄기류 박람회' 개최.
<p>5. 열매/과일류 활성화 방안</p> <ul style="list-style-type: none"> • 취미와 요리 관심층 타겟팅 <ul style="list-style-type: none"> • 열매/과일을 활용한 DIY 잼 만들기, 디저트 레시피 클래스 운영. • 열매/과일류를 활용한 홈카페 레시피를 SNS에 공유하고 챌린지 형태로 확산. • 학생 및 젊은 층 타겟팅 <ul style="list-style-type: none"> • 열매/과일을 활용한 건강 간식 상품 개발, 특히 '스무디용 과일 세트' 판매. • 간편 포장된 '하루 과일' 제품 라인 출시. • 봄철 마케팅 <ul style="list-style-type: none"> • 봄/초여름 시즌 캠페인으로 신선한 과일을 활용한 "제철 과일 축제" 진행. • 지역 특산 열매/과일의 산지 직송 프로모션. 	<p>공통 활성화 전략</p> <ul style="list-style-type: none"> • 봄철 집중 마케팅 <ul style="list-style-type: none"> • 봄철 및 초여름 시즌 관심도를 반영하여 모든 임신물에 대한 프로모션을 이 시기에 집중. • 가족 단위 이벤트 <ul style="list-style-type: none"> • 자녀가 있는 소비자층을 겨냥해 '가족과 함께하는 요리 클래스' 또는 임신물 체험 행사 진행. • 온라인 플랫폼 활용 <ul style="list-style-type: none"> • 다양한 레시피를 제공하는 웹사이트 및 앱 구축. • 구매자의 후기와 요리 사진 공유 이벤트로 참여를 유도.

그림 13 생성형AI 활성화 방안 생성 결과

그림 13은 생성형 AI에 분석 인사이트를 담은 프롬프트를 입력하여 얻은 임신물 종류별 활성화 방안 결과이다. 각 임신물 종류별로 관심사, 주 연령층, 계절 등의 여러 인사이트를 맞춤형으로 충분히 고려하여 현실적이고 구체적인 활성화 방안을 안내해주고 있음을 확인할 수 있다.

임산물 종류	활성화 방안	근거
건과류	<ul style="list-style-type: none"> * 학생들의 관심을 반영하여 건과류를 활용한 간편 간식 제품 개발 * SNS플랫폼을 통한 "건과류로 만드는 5분 간식" 콘텐츠 제작 * 건과류의 영양적 가치를 강조하여 20~40대가 주로 사용하는 피트니스 앱이나 헬스 건강 매체에 광고 송출 	<p>[insight2] 학생들을 주로 타겟</p> <p>[insight5] 임신물을 이용한 건강에 관심있는 20, 40대를 주로 타겟</p>

나물류	<ul style="list-style-type: none"> * 전령층을 대상으로 나물 요리법 강좌 및 유튜브 콘텐츠 업로드 * 주부층을 겨냥해 홈쇼핑에서 간단히 조리 가능한 '손질된 나물 키트', '봄철 인기 나물 세트' 기획 판매 및 QR코드로 연결되는 레시피 제공 * 자연 체험 프로그램과 연계해 가족 단위의 나물 채집 행사 기획 	[insight1] TV매체,요리/레시피와 취미에 관심있는 사람들이 관심 [insight2] 주부들을 주로 타겟 [insight3] 봄철과 초여름을 타겟 [insight5] 요리/레시피에 관심있는 전 연령층을 타겟
버섯류	<ul style="list-style-type: none"> * 면역력 강화와 건강을 주제로 한 버섯 관련 마케팅 진행 * 학생들을 위한 '버섯 키우기 키트'를 개발하여 판매 	[insight1] 건강과 판매정보에 관심있는 사람들을 타겟 [insight2] 학생들을 주로 타겟 [insight5] 임산물을 이용한 건강에 관심있는 20, 40대를 주로 타겟
뿌리/줄기류	<ul style="list-style-type: none"> * 뿌리/줄기류를 활용한 건강주스, 간단 요리법을 TV 매체와 유튜브 채널에서 홍보 * 자녀가 있는 30-50대를 겨냥해 건강 효능을 강조 * 봄철 농가와 협력해 다양한 요리법을 박람회 형태로 개최 	[insight1] TV매체에 관심 있는 사람들을 주로 타겟 [insight3] 봄철과 초여름을 타겟 [insight4] 결혼한 사람 중, 자녀가 있는 사람들을 주로 타겟
열매/과일류	<ul style="list-style-type: none"> * 열매/과일을 활용한 DIY 잼 만들기 SNS(유튜브, 인스타그램 릴스)에 홈카페 레시피 공유 * 간편 포장된 '하루 과일' 제품 라인 출시 	[insight1] 취미와 판매정보에 관심 있는 사람들을 주로 타겟 [insight5] 요리/레시피에 관심있는 전 연령층을 타겟

표 5 생성형AI 활성화 방안 생성 결과

이를 토대로, 본 프로젝트에서 제안하고자 하는 임산물 종류별 활성화 방안은 표 5와 같다.

3) 임업통계 등 데이터 적용 방안

필수) 산림통계시스템 '임산물 생산비조사' 데이터: 해당 데이터는 임산물별(밤, 대추, 호두 등) 생산비를 나타낸 데이터로, 그중 11가지 대표 임산물에 대한 2023년 임산물별 1ha당 소득분석(단위: 평가액(원)) 데이터의 순수익 컬럼을 사용한다. 이를 통해 임산물의 순수익 정도 및 대부분의 임산물의 순수익이 적자라는 현재의 문제점을 잘 파악할 수 있다.

필수) 산림청 산림임업통계플랫폼의 인포그래픽 '2023년 임업경영실태조사 주요결과' 데이터: 해당 데이터는 2023년의 임업경영실태조사 주요 결과를 나타내며, 인포그래픽 데이터를 엑셀로 변환하여 활용하였다. 임산물구분명, 월별 조회수, 판매수입, 투입비용, 소득 동향, 생산량 동향, 업종 전망 등 주요 데이터를 추출한 뒤, 22년과 23년의 관심도 점수를 산출하였다. 또한, 관심도와 관심 지표를 시각적으로 표현하여 업종 간 차이를 확인했다.

필수) 한국임업진흥원 '임가소득의 구성요소별 증감률' 데이터: 한국임업진흥원의 '통계 정보'에서 확인할 수 있는 이 데이터는 임가 경제 및 임업경영실태를 나타낸 데이터이다. 임가소득, 경상소득, 비경상소득, 임업의존도가 기록되어 있으며, 임업 소득이 하락세라는 현 상황의 문제점을 입증하는 용도로 사용한다.

선택) 산림 빅데이터 플랫폼의 '2019~2020년 성/연령/지역 등 이용자 분류별 6대 임산물 관심도 정보' 데이터: 감, 대추, 더덕, 밤, 표고버섯, 호두에 대한 연령, 성별, 지역별 관심도를 분석한 자료로,

2019년부터 2020년까지 수집된 데이터이다. 이 데이터는 18,667개의 행과 연령 그룹, 관심도(빈도수), 추출 단어(임산물 종류)와 같은 주요 변수로 구성되어 있다. 이를 통해 연령대별 임산물 선호도를 분석하고, 특정 임산물에 대한 소비 경향을 도출한다.

선택) 산림 빅데이터 플랫폼의 '농산물 소비인구 정보' 데이터: 임산물을 포함한 더 넓은 범위의 농산물 소비인구 데이터로, 임산물을 구매한 인구와 그렇지 않은 인구를 분류하는 모델을 구축하는 데이터로 활용한다. 2020년에 1년간 수집된 데이터이고, 39,127개의 행과 22개의 변수로 구성되어 있다. 이용자성별코드, 연령, 이용자직업명, 가구소득그룹명, 정보획득사이트URL, 제품내용 등의 정보를 포함한다.

선택) 산림 빅데이터 플랫폼의 '임산물 관심인구 정보' 데이터: 검색/블로그/뉴스기사 등의 사이트에서 임산물에 관심을 보인 인구에 대한 정보를 담은 데이터로, 1단계에서는 임산물별 사람들의 관심 정도 및 특성 분석을 위해 활용하며, 2단계에서는 데이터의 절반은 임산물 구매 분류 모델의 학습데이터로 사용되고, 절반은 구축된 분류 모델의 테스트 데이터 및 관심사 클러스터링을 위한 데이터로 활용한다. 또한, 분류 모델과 관심사 클러스터링의 결과를 종합해 임산물을 활성화하기 위한 방안을 제공하는 데이터로 활용한다. 2019년 12월부터 2020년까지 수집된 데이터이고, 18,667개의 행과 20개의 변수로 구성되어 있다. 위의 데이터와 마찬가지로 이용자성별코드, 연령, 이용자직업명, 가구소득그룹명, 정보획득사이트URL, 제품내용 등의 정보를 포함한다.

4) 기대효과

기술적 측면

기술적 측면의 기대효과는 다음과 같이 요약할 수 있다. 첫째, 우리가 쉽게 접할 수 있는 검색, 블로그, 뉴스 기사 등에서 임산물에 관심을 보이는 사람들이 실제로 임산물을 구매할 잠재 고객인지 예측할 수 있는 모델을 개발함으로써, 추후에도 잠재 고객을 빠르고 정확하게 파악할 수 있는 도구로 활용할 수 있을 것으로 기대한다. 둘째, 임산물 관심 인구의 관심사를 클러스터링하여 분류함으로써, 각 임산물의 잠재 고객들이 대표적으로 관심 있는 분야를 추출할 수 있게 되고, 이는 임산물 분류별 활성화 방안을 수립하는데 있어 핵심 요소로 활용이 가능하다. 셋째, 분석을 통해 직접 도출한 여러 인사이트를 반영해 생성한 프롬프트와 생성형 AI의 기술을 결합해 임산물별 최적의 활성화 방안을 제시하는 프로세스를 제공함으로써, 임산물 이외의 여러 임업 관련 분야에서도 확장할 수 있는 가능성을 지닌다.

산업적 측면

산업적 측면의 기대효과는 다음과 같이 정리할 수 있다. 첫째, 각 임산물 종류별로 잠재 고객의 관심사와 특성을 파악하고, 타겟 고객별 맞춤형 활성화 방안을 마련함으로써 소비자의 관심을 효과적으로 유도하고, 임산물에 대한 긍정적인 인식을 제고할 것으로 기대된다. 둘째, 생산량에 비해 판매량이 저조했던 비인기 임산물의 경우에도, 효과적인 활성화 방안을 통해 해당 임산물에 대한 소비자의 관심과 수요를 높일 수 있다. 이를 통해 생산 과정에서 발생하는 불필요한 낭비를 줄이고, 버려지는 임산물의 양을 감소시킴으로써 자원의 효율적 활용이 가능해질 것이다. 셋째, 임산물 소비가 증가함에 따라 임업 종사자들에게 더 많은 경제적 기회가 제공될 것으로 기대된다. 이를 통해 기존의 수익 구조를 보완하고 추가적인 수익원을 확보할 수 있을 뿐만 아니라, 관련 산업 전반에도 긍정적인 영향을 미칠 것이다.

정책적 측면

정책적 측면의 기대효과는 다음과 같다. 첫째, 각 임산물 종류별로 수집된 데이터 기반으로 세분화되

면서도 현실적인 답변을 할 수 있도록하는 프롬프트 지시를 통해 실효성 있는 임업 관련 정책안을 도출할 수 있다. 둘째, 본 프로젝트의 인사이트와 생성형AI의 문장 생성 능력을 결합함으로써 기존에 생각하지 못한 창의적인 활성화 정책을 수립할 수 있을 것으로 기대한다. 셋째, 기존에는 생산자, 즉 공급 측면의 임산물 정책이 많은 반면[정책1], 본 프로젝트를 통해 임산물 구매자를 증가시킬 수 있는 임산물의 수요 관점에서의 구체적인 활성화 방안을 마련함으로써 공급 뿐만 아니라 기존의 한계점이었던 수요 해결을 통해 근본적인 문제 해결 방안을 마련한다.

4) 레퍼런스

- [1] "간화군산림조합, '임산물 소비촉진' 행사 마련", 시민일보, 문찬식, 2023.10.20., <https://www.siminilbo.co.kr/news/newsview.php?ncode=1160304821864891>
- [2] "임산물 판매망 다변화...농가소득 높여", 농민신문, 양석훈, 2022.12.08., <https://www.nongmin.com/368122>
- [3] Doe, J., & Smith, A. (2023). A Study on Neural Networks. *arXiv*. <https://arxiv.org/abs/2312.16171>
- [4] 변승연, & 석현덕 (2019). 산채(산나물)에 대한 소비자 의향 및 수요 분석: 소비자 의향 조사와 소셜 빅데이터 분석을 통하여.
- [5] Bsharat, S.M., Myrzakhan, A., & Shen, Z. (2023). Principled Instructions Are All You Need for Questioning LLaMA-1/2, GPT-3.5/4. ArXiv, [abs/2312.16171](https://arxiv.org/abs/2312.16171).
- [6] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- [7] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. <https://doi.org/10.1145/2939672.2939785>
- [8] Duan, T., Anand, A., Gong, Z., Shabat, E., Chang, S., & Ng, A. Y. (2020). NGBoost: Natural gradient boosting for probabilistic prediction. Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI). <https://arxiv.org/abs/1910.03225>
- [9] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533–536. <https://doi.org/10.1038/323533a0>
- [10] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research, 3(Jan), 993–1022. <https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf>
- [11] Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. ArXiv preprint [arXiv:2203.05794](https://arxiv.org/abs/2203.05794). <https://arxiv.org/abs/2203.05794>.
- [12] Ukairia777. (2022). KoBERTopic: Korean adaptation of BERTopic for neural topic modeling. GitHub repository. Retrieved November 21, 2024, from <https://github.com/ukairia777/KoBERTopic>
- [13] OpenAI. (2023). ChatGPT. <https://chatgpt.com/>.