

# [기획서] 데이터 기반 유튜브 썸네일 성과 예측 및 최적화 솔루션

## 1. 프로젝트 개요 (Project Overview)

- 프로젝트명: SceneClick: 회귀분석 기반의 유튜브 썸네일 AI 진단 솔루션
- 배경: 유튜브 레드오션 시대, 영상의 품질보다 중요한 것은 '클릭을 유도하는 썸네일'임.
  - 기존 디자인 툴은 '심미성'에만 집중할 뿐, 실제 '조회수(성과)'와 연결되는 '팔리는 디자인 법칙'을 제시하지 못함.
  - 목표: 2024년 해외여행 브이로그 유튜브 썸네일 데이터를 분석하여 조회수와 썸네일 요소 간의 상관관계(회귀식)를 도출.
  - 기존 썸네일을 분석하여 문제점을 알려주고, 조회수 상승을 위한 구체적인 개선안 생성.

## 2. 핵심 타겟 (Target Audience)

- 메인 타겟: 구독자 수 대비 조회수 효율을 극대화하고 싶은 유튜버.
- 니즈: "내 영상이 재미없어서 안 보는 걸까, 썸네일 때문에 클릭을 안 하는 걸까?"에 대한 객관적 해답 요구.

## 3. 데이터 수집 및 가공 전략 (Data Strategy)

본 프로젝트의 핵심 경쟁력은 '정교한 데이터 설계'에 있음.

### A. 데이터셋 구축 범위

- 기간: 2024년 1월 1일 ~ 2024년 12월 31일 (최근 1년 데이터로 트렌드 반영 및 계절성 통제)
- 대상: 특정 카테고리(예: 해외여행 브이로그, 엔터, 정보성 등)의 유튜버 영상 데이터.

### B. 변수 설계 (Feature Engineering)

단순 수집을 넘어, 비정형 데이터(이미지)를 정량적 데이터(숫자)로 변환함.

구분	변수명	설명 및 추출 방식	비고
종속 변수(Y)	예상 조회수 효율(%)	'영상 조회수/구독자 수 * 100'	채널 체급을 제외하고 순수 썸네일 효율 측정
독립 변수(X)	시각적 요소 구성 요소	밝기, 대비, 채도 얼굴 면적 비율, 텍스트 면적 비율	썸네일 효율 측정 이미지 기초 분석
통제 변수	최신성	13 - 게시월 (최신일수록 높은 점수)	누적 시간의 불공정성 보정

## 4. 핵심 기술 및 알고리즘 (Methodology)

### Step 1. 특징 추출 (Feature Extraction)

- 사용자가 기존 썸네일을 업로드하면 AI가 이미지를 분석.
- ex) "얼굴이 전체의 30%를 차지함" 등의 정보를 수치 데이터로 변환.

### Step 2. 회귀분석 (Regression Analysis)

- 수집된 데이터를 바탕으로 다중 회귀분석(Multiple Regression)수행.
- 도출 목표: 예상 조회수 효율(%) (Y)에 가장 큰 영향을 미치는 '핵심 변수(X)' 얼굴.
- 결과: 화면에서 얼굴이 1% 커질 때마다 예상 조회수 효율은 40.8%p하락

### Step 3. 솔루션 제공 (Evaluation & Feedback)

- 예측 점수: 회귀방정식에 사용자 데이터를 대입하여 예상 성과 지수 산출.
- 진단 리포트:
  - "현재 점수: 65점 (평균 이하)"
  - "솔루션: 현재 텍스트 크기가 너무 작습니다. 텍스트 면적을 20% 이상으로 키우면 예상 점수가 80점으로 상승합니다."

## 5. 기대 효과 및 차별점 (Differentiation)

구분	기존 서비스	SceneClick
기준	"예쁜가?" (주관적 디자인)	"클릭을 부르는가?" (데이터 기반)
근거	디자이너의 감	2024년 실데이터 회귀분석
결과물	이미지 편집 도구 제공	성과 예측 및 개선된 결과물 생성

## 6. 향후 발전 계획 (Roadmap)

- 데이터셋 확장을 통한 카테고리별(먹방, 게임, 브이로그) 맞춤형 회귀 모델 세분화.
- 생성형 AI를 연동하여 진단 내용을 바탕으로 "최적화된 썸네일 자동 생성"기능까지 확장.