Azure Document Intelligence -Receipts vs Invoices 모델 비교 실 험 보고서

1. 📌 목적

- **주 사용 모델**: Prebuilt Receipts (영수증)
- 문제 상황: 일부 인보이스 데이터가 포함되어 있음
- 고려 사항:
 - Receipts 모델로 인보이스까지 커버할 수 있는지?
 - Invoice 모델을 병행 사용해야 할 만큼 성능 차이가 있는지?

2. 🔪 실험 환경

- 모델 종류
 - Prebuilt Receipts
 - Prebuilt Invoices
- 테스트 데이터셋: https://github.com/mouadhamri/invoice_dataset
- **데이터 수**: 총 10건
- **언어**: 영어
- 실행 방식: Azure Studio

3. 영수증 모델로 실험 결과

파일: FACTU2017070047

실제 문서 유형: Invoice (프랑스어, 세금 포함 인보이스)

예상된 분석 목표: Receipt 모델이 인보이스 문서를 어떻게 처리하는가

📤 Receipts 모델 분석 결과 요약

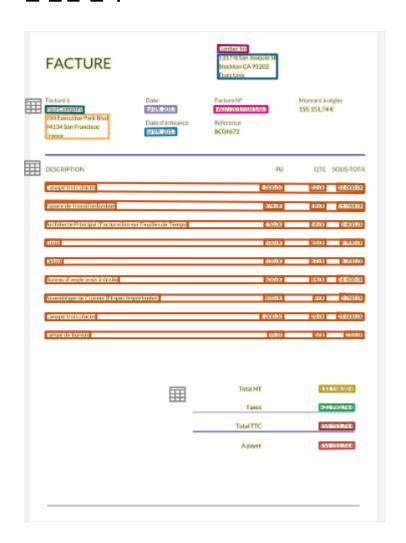
FACTURE		1337 N San Joaquin St Stockton CA 95202 États Unis	
Facturé à Your Company 250 Executive Park Blvd 94034 San Francisco France	Date MC9, 2017 Date d'écheance MC9, 2017	Facture N° FA07/2017/041461 Reference BC04672	Monrant à régler 155 151.74 €
DESCRIPTION		PU	QTE SOUS-TO
Canapé trois places		100000	32,00 32,000)
Espace de travail individue		876,00	18,00 15 768,
Architecte Principal (Facturations	ur Feuilles de Temps	150.00	16,00 2400)
Tiroir		100.00	69,00 6 900)
Depot		100.00	67,00 6 700)
Bureau d'angle assis à droite		800.00	44,00 26 400)
Assemblage de Cuisine (Etapes Im	portantes)	420,00	9,00 3 780,
Canapé trois places		100000	39,00 39,000)
Lampe de bureau		35,00	6,00 210.0
		Total HT	133158.00€
	_	Taxes Total TTC	155150740
	_	Apayer	155 151,74€
		Apper	155 152,/4€

항목	결과	비고
DocType	receipt.retailMeal	❤️ 오분류 (실제는 인보이스)
MerchantName	Lumber Inc (97.9%)	✓ 정확
MerchantAddress	1337 N San Joaquin St	✓ 주소 파싱도 잘됨
TransactionDate	2017-07-09 (98.7%)	✓ 정확
ReceiptType	Meal (98.0%)	★ 오분류
Subtotal	EUR 133,158	✓ 추출됨
TotalTax	EUR 21,993.74	✓ 추출됨
Total	EUR 155,151.74	✓ 정확
Items	9개 추출	✓ 개수 정확 (내용 확인 필요)

🧠 1차 평가

평가 기준	내용
문서 유형 분류	★ 인보이스를 영수증으로 오분류함
핵심 필드 추출 정확도	☑ 금액, 날짜, 상호명 등은 잘 추출됨
비정상 필드 추출	ReceiptType = Meal → 명백한 오분류
사용 가능성	문서 유형에 따른 조건 분기 없으면 오류 가능성 있음

4. 인보이스 모델 실험 결과



☑ Receipts vs Invoices 모델 비교

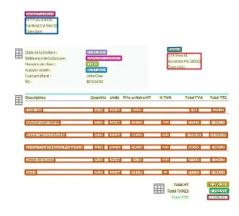
필드	Receipts 모델	Invoices 모델	차이 요약
DocType	receipt.retailMeal	invoice 🗸	오분류 vs 정확

VendorName / MerchantName	Lumber Inc	Lumber Inc 🗸	동일
VendorAddress / MerchantAddress	Stockton, CA 주소	동일 🗸	동일
Invoiceld	🗶 없음	FA07/2017/041461 🗸	Receipts는 미지원
InvoiceDate / TransactionDate	2017-07-09	2017-07-09	동일
Subtotal	EUR 133158 🗸	EUR 133158 🗸	동일
TotalTax	EUR 21993.74 🗸	EUR 21993.74 🗸	동일
Total	EUR 155151.74 🗸	EUR 155151.74 🗸	동일
AmountDue	່✗없음	EUR 155151.74	Receipts 모델에는 미포함
CustomerName	່✗없음	YourCompany	Invoices 모델만 추 출
BillingAddress	່✗없음	250 Executive Park Blvd	Invoices 모델만 추 출

☑ 평가 요약

항목	평가
정확한 DocType 분류	Invoice 모델이 정확함
필드 커버리지	Invoice 모델이 훨씬 넓음
금액/날짜 추출 정확도	두 모델 모두 정확
고객 정보 및 송장 ID	Invoice 모델만 지원
실제 사용 시 적합도	Invoice 모델 우세

비교 문서: FACTU2017050053.png



<인보이스 모델 결과>



<영수증 모델 결과>

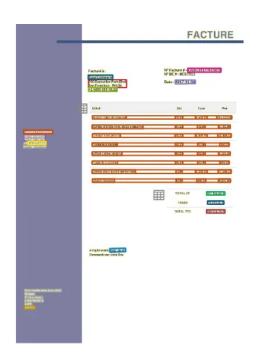
문서 유형: Invoice (프랑스어, 고객/공급자 정보 포함)

☑ Receipts vs Invoices 모델 비교

필드	Receipts 모델	Invoices 모델	차이 요약	
----	-------------	-------------	-------	--

DocType	receipt.retailMeal	invoice 🗸	오분류 vs 정확
MerchantName / VendorName	Gemini Furniture	Gemini Furniture	동일
MerchantAddress / VendorAddress	Fairfield CA 주소	Fairfield CA 주소	동일
TransactionDate / InvoiceDate	2017-07-05 🗸	2017-07-05	동일
InvoiceId	່✗없음	FA05/2017/042475	Receipts 모델은 미 지원
Subtotal	EUR 45295.50 🗸	EUR 45295.50 🗸	동일
TotalTax	EUR 7752.55 🗸	EUR 7752.55 🗸	동일
Total	EUR 53048.05 🗸	EUR 53048.05 🗸	동일
CustomerName	🗙 없음	Biotech <	Invoices만 지원
CustomerAddress	🗙 없음	215 Vine St 🗸	Invoices만 지원
CustomerId	່✗없음	CL002 🗸	Invoices만 지원

마 비교 문서: FACTU2017040057.png



<인보이스 모델 결과>



<영수증 모델 결과>

문서 유형: Invoice (고객/공급자 정보, 전화번호, Swift 포함)

📊 Receipts vs Invoices 모델 비교

필드	Receipts 모델	Invoices 모델	차이 요약
DocType	receipt.retailMeal 🗶	invoice <	오분류 vs 정확
MerchantName / VendorName	Gemini Furniture 🗸	Gemini Furniture 🗸	동일
MerchantAddress / VendorAddress	San Francisco / Fairfield 주소 🗸	동일 🗸	동일
TransactionDate / InvoiceDate	2017-11-04 🗸	2017-11-04 🗸	동일
Invoiceld	່★없음	FA04/2017/065819	Invoices 전용
Subtotal	EUR 374186.5 🗸	EUR 374186.5 🗸	동일
TotalTax	EUR 74373.2 🗸	EUR 74373.2 🗸	동일
Total	EUR 448559.7 🗸	EUR 448559.7 🗸	동일
CustomerName	່★없음	YourCompany <a>V	Receipts 모델 누락

CustomerAddress	່✗없음	250 Executive Park Blvd	Invoices만 지원
PhoneNumber	▼ 추출됨	✓ 추출됨	양쪽 모두 존재 (신뢰 도는 낮음)
PaymentDetails (SWIFT)	★ 없음	SF7575 V	Invoices 전용 필드

☑ 평가 요약

항목	평가
문서 유형 분류	Invoice 모델이 정확, Receipts는 오분류
핵심 금액 필드	동일하게 잘 추출
고객/송장 관련 필드	Invoice 모델만 추출
결제 수단 (SWIFT)	Invoice만 지원
Receipts 모델 한계	인보이스 문서에 대해 비즈니스 필드 대부분 누락

🔽 최종 결론 요약

- DocType 분류 정확도:
 - Receipts 모델은 실험한 3건 모두에서 문서 유형을 잘못 분류했음 (receipt.retailMeal 로 오인식)
 - Invoice 모델은 모두 정확하게 invoice 로 분류함
- 금액 관련 필드 (Subtotal, TotalTax, Total)
 - Receipts 모델과 Invoice 모델 모두 동일하게 정확히 추출됨
 - 。 금액 필드만 필요한 경우 Receipts 모델도 충분히 활용 가능함
- Invoice 전용 필드
 - Invoice 모델만 추출 가능한 필드 존재:
 - InvoiceId
 - CustomerName
 - BillingAddress , SWIFT , CustomerId 등
 - ∘ 해당 필드가 필요한 경우 Receipts 모델로는 불충분



💡 실무 적용 가이드

사용 목적	추천 모델	사유
금액 추출 (정산용, 분석 용)	Receipts 모델 단독 사용 가능	금액 필드는 정확하게 추출됨, 후처리 없이 사용 가능
문서 분류 및 송장 정보 추출	Invoices 모델 사용 필요	고객 정보, 송장 ID 등은 Receipts 모델에서 누락됨
혼합 문서 처리	문서 분류 후 분기 처리 추 천	영수증/인보이스 혼합 상황에서는 모델 분기 가 안정적