Alcohol by Volumen (ABV)

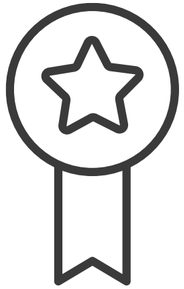
Prediction Platform

Business Analytics & Insights
Max Causso – 301455365

Executive Summary

Objective

- Build an ABV (alcohol%) predictor to adjust recipes before brewing, reducing rework and label risk in Canada.



Business Value

- Fewer out-of-tolerance batches, faster releases, and protected excise margin.

Solution

- Azure cloud architecture.
- Test multiple algorithms and choose the best accuracy + stability + explainability.

Business context & Opportunity

Labelling

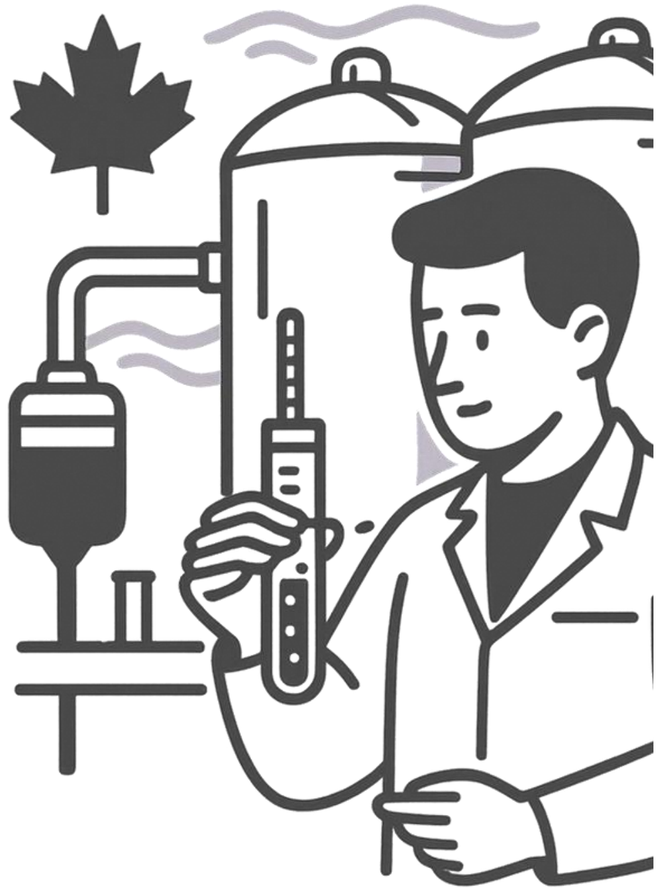
- Canadian Food and Drug Regulations require the % alcohol declaration when ≥1.1% ABV.

Provincial tolerances:

- **Ontario (LCBO):** Use LCBO's "Actual vs Declared Alcohol Content – Guideline."
- **Nova Scotia (NSLC):** Explicit bands (e.g., 4.1–5.5% ABV: ±0.5% is adherence).

Operational risk & cost

- Early ABV prediction reduces rework, relabeling, QC holds, and scrap, protecting margin and speeding releases.



Sources: CFIA – Labelling requirements for alcoholic beverages; LCBO QA – Actual vs Declared Alcohol Content – Guideline; NSLC – ABV Factsheet

Approach overview

1

Data Preparation

- Cleaned data, grouped styles.
- Engineered OG–FG (attenuation) and ratios.

2

Feature Selection

- Compared linear (Full/Ridge), tree/ensemble, and NN
- Balanced accuracy vs explainability.

3

Model testing & evaluation

- Hold-out metrics (R^2, MAE).
- Controlled overfitting.
- Ridge kept as glass-box baseline.

4

Deployment in Cloud

- Azure API + Power BI dashboard.
- Versioned model & audit log.

Model Comparison

Full Regression Companion (glass-box)

- R^2 0.869
- Use: Explainability/training, audits, coaching
- Not used as the operational predictor

Random Forest

- R^2 0.995 – MAE 0.055% ABV
- Captures non-linear recipe effects
- Stable across styles and low CV variance
- Use: Operational predictor (pre-brew)

Final choice: Random Forest

Accuracy with regulatory headroom

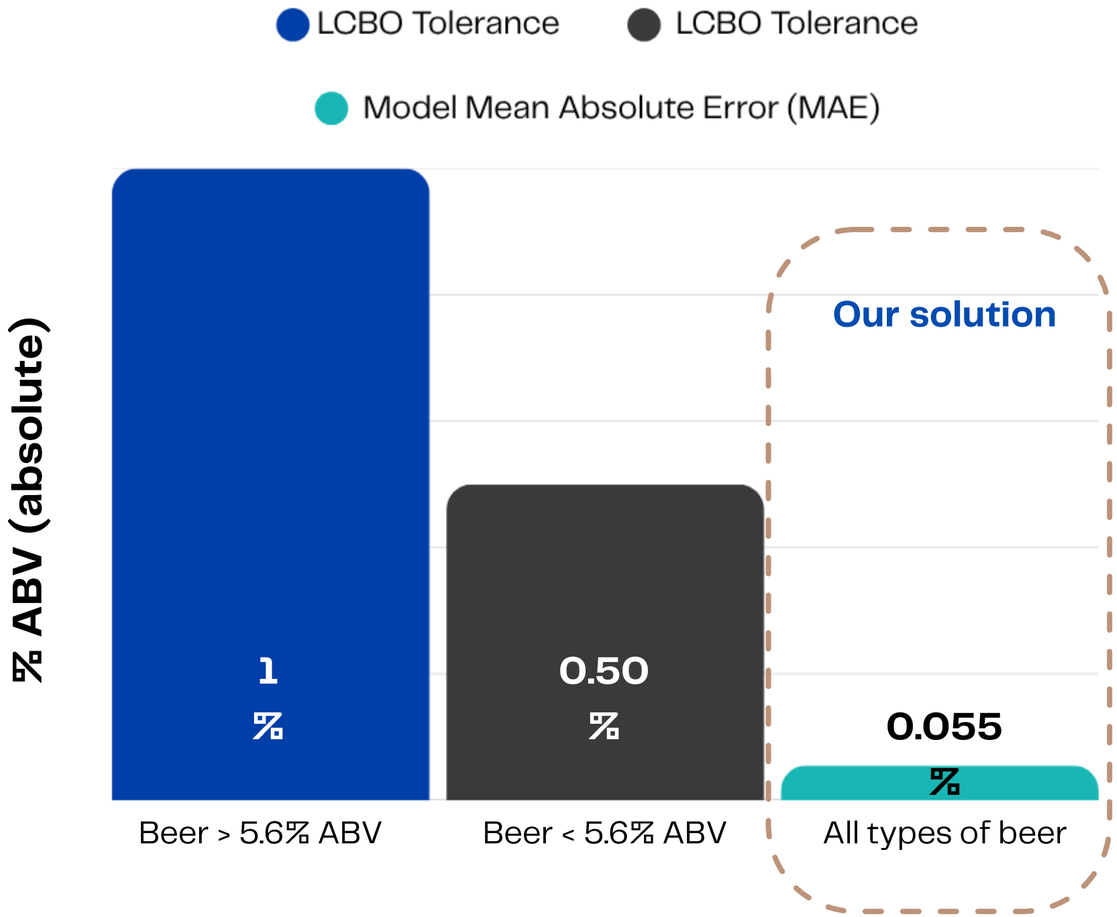
- R^2 0.995, MAE 0.055% ABV — ≈9× smaller than LCBO one-side tolerance (0.50) for beers <5.6% ABV.

Lower label/compliance risk

- Fewer lots near tolerance edges → fewer holds, retests, or relabels. Predictions versioned in Azure SQL for QA/audits.

Operational fit & speed

- Batch scoring in Python; ADF orchestration; Azure SQL serving; Power BI visibility. Minutes, not hours.



Sources: LCBO QA – Product Packaging Standards & Guidelines for Chemical Analysis (p. 52, "Actual vs Declared Alcohol Content – Guideline"); NSLC – ABV Factsheet

Top drivers of ABV

Attenuation (OG – FG) dominates

- Quick lever: yeast health/pitch, ferm temp/time, mash temp.
- Direction: lower FG : ↑ABV; higher FG : ↓ABV (ABV ≈ 131.25 × ΔSG)

Original Gravity (OG) sets the ceiling

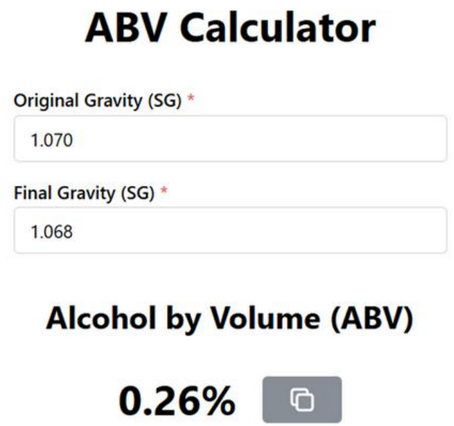
- Quick lever: grain bill concentration, adjuncts, boil concentration.
- Direction: higher OG: ↑ABV if FG doesn't rise much (diminishing returns at high OG)

Efficiency × OG (real-world limiter)

- Quick lever: lauter performance, mill gap, mash schedule, sparge discipline.
- Direction: better efficiency at target OG: ↑ABV; at extremes returns flatten

Real-life examples

How tweaks change ABV



Recipe level (OG & FG)

- +0.25% ABV → target −0.002 SG in FG (2 points).
- −0.25% ABV → allow +0.002 SG in FG.
- +0.6–1.3% ABV → raise OG by +0.005–0.010 SG and protect attenuation.

Compliance check (ON/NS)

- Compare to LCBO/NSLC band.
- House guardrail: escalate to QA if margin ≤0.2% to nearest limit.
- Example: 5.34% vs upper 5.50 → margin 0.16 → Escalate .

Sources: ABVCalculator.ca; CFIA — Labelling requirements for alcoholic beverages; LCBO QA — Product Packaging Standards & Guidelines for Chemical Analysis,

Pattern by beer style

ABV range

Use BJCP style ranges as QA guardrails

- Lag any predicted ABV that sits outside the style's window before brewing. (e.g., many international lagers around ~4–5% ABV; styles like imperial stouts much higher — see per-style sheets.)

Style	BJCP code	ABV range (%)
International Pale Lager	2A	4.5 – 6.0
American IPA	14B	5.5 – 7.5
Belgian Tripel	26C	7.5 – 9.5
Imperial Stout	20C	8.0 – 12.0

Sources: ABVCalculator.ca; CFIA — Labelling requirements for alcoholic

Cloud-based prediction platform

Inputs

- OG, expected FG (target attenuation), brewhouse efficiency, IBU, style.

Outputs

- Predicted ABV + confidence band + green/amber/red tolerance light.

Integration

- Power BI dashboard for QA approvals.
- Versioned model service.
- Audit log of predictions.

Business value for breweries

Quality control — right-first-time

- Pre-brew flags using LCBO/NSLC guardrails
- Fewer rework/hold decisions, faster QA sign-off
- Upstream ABV decisions (before wort is made)

Variability reduction – stable FG, better yield

- Standardize pitch rate / O_2 / fermentation temperature.
- Protect attenuation → tighter FG, less ABV drift.
- Reduce over-strength giveaway

Compliance & margin protection

- On-label across LCBO/NSLC; fewer relabel cycles
- Excise risk: avoid crossing CRA tiers
- Audit-ready logs in Azure SQL & Power BI

Recommendations & Next steps

Adopt

Random Forest: for day-to-day; keep Full (Ridge) Regression for explainability/audits.

Instrumentation

Capture OG & FG per batch; correct for temperature; document yeast pitch rate, fermentation T°, and pH in a standard log.

Dashboards

Add BJCP guardrails + province-specific tolerance lights (e.g., NSLC table; LCBO guideline page).

Pilot

(30–60 days, 2 brands): KPIs = % batches out-of-tolerance, rework hours, days-to-release; scale on success.

Thanks