

# **Project Report: Insurance Quotation Automation Using Rasa Chatbot**

**Dated: 11th November 2024**

## **Quotation Quokkas Members**

Patricia Ngari

Paullet Ndeda

Brandon Mwangi

Gloria Ngure

## TABLE OF CONTENTS

<b>Business Understanding.....</b>	<b>1</b>
Business Overview.....	1
Business Objective.....	2
Business Success Criteria.....	2
Requirements.....	2
Assumptions and Constraints.....	2
Data Mining Goals.....	3
Data Mining Success Criteria.....	3
<b>Data Understanding.....</b>	<b>3</b>
Data Understanding Overview.....	3
Data Description.....	3
Verifying Data Quality.....	4
Data Mining Success Criteria.....	4
Data Mining Goals.....	4
<b>Data Preparation.....</b>	<b>4</b>
Loading Data.....	4
Cleaning Data.....	5
<b>Analysis.....</b>	<b>5</b>
<b>Recommendations.....</b>	<b>5</b>

# Business Understanding

## Business Overview

Our project focused on developing a chatbot to generate Professional Indemnity Insurance (PII) quotations, specifically catering to a wide range of business professionals. The chatbot collected key information from users, including details about business profession, reinsured, broker, insured names, staffing information, and indemnity amount before generating a customized insurance quote based on this input.

## Business Objective

The project's primary goal is to automate the insurance quotation process, ensuring that clients receive accurate and timely quotes with minimal manual intervention. The chatbot provided an efficient, scalable solution that generated quotes based on user input, streamlining the process for both the insurer and the client.

## Business Success Criteria

We defined the success of the project by the following criteria:

- **Accurate and reliable quote calculations** - ensure that all fees, premiums, and indemnities were calculated correctly.
- **User satisfaction** - ensure that user experience is intuitive, fast, and error-free.
- **Scalability** - ensure that the system is capable of handling additional business professions and insurance parameters in the future.
- **Compliance** - ensure that the premium and indemnity calculations comply with industry standards.

## Requirements

Our chatbot required the following capabilities to function effectively:

1. **Extract, validate, and analyze user input** - process details such as business profession, names of the insured, and staffing levels.
2. **Calculate insurance premiums, fees, and indemnities** - based on predefined business logic and professional guidelines.
3. **Generate quotations** - based on the input data, generating the quote for review and distribution.

## Assumptions and Constraints

We assumed that users would provide all necessary data inputs, including staffing numbers, indemnity amounts, and optional coverage selections. We encountered the following constraints:

- **Data completeness** - the system assumed that all required inputs were given by the user.
- **Limitation on certain high-risk professions** - the chatbot did not support certain high-risk professions, which could be flagged and excluded from receiving a quote.
- **Data security and calculation accuracy** - we ensured that all calculations were based on accurate, secure data to prevent any discrepancies.

## Data Mining Goals

Our data mining goals included:

- **Mapping business data to fee calculations** - applying structured logic to derive accurate premiums, staff fees, and other values.
- **Deriving actionable insights** - based on the data provided by users, we extracted key insights to enhance the premium calculation.
- **Validating quote results** - ensuring that the results of the quote met predefined success criteria.

## Data Mining Success Criteria

We determined the success of our data mining process based on the following:

- **Accurate data handling** - ensuring that all collected data was correctly mapped to the premium and indemnity calculation logic.
- **Effective input validation** - ensuring no erroneous data led to inaccurate results.
- **Error-free premium calculations** - all calculated premiums needed to be valid based on business rules.

# Data Understanding

## Data Understanding Overview

The data used in our chatbot was primarily user-provided input, which included various fields essential for generating an accurate insurance quote. This included information on the business profession, names of the reinsured, broker, and insured entities, staffing details (e.g., number of staff members in various categories), indemnity amounts, and optional extensions.

## Data Description

The primary data fields we collected through the chatbot interface included:

- **Business Profession** - the type of business applying for insurance (e.g., architect, accountant); impacts risk assessment.
- **Entity Names** - names of the reinsured, broker, and insured parties; used for legal identification.
- **Staffing Information**: These metrics inform workforce composition and potential operational risk.
  - **Partners/Principals** - number of business partners or principal members.
  - **Qualified Assistants** - count of formally qualified assistants.
  - **Unqualified Assistants** - count of assistants without formal qualifications.
  - **Other Staff** - total other staff members.
- **Indemnity Amount** - requested insurance coverage limit; a key factor in premium calculation.

## Verifying Data Quality

We verified the data for completeness and accuracy:

1. **Complete Input** - the chatbot ensured that users provided all required fields (e.g., staffing numbers, indemnity amounts).
2. **Accuracy** - we validated numeric values to ensure they were within valid ranges.
3. **Handling Missing or Erroneous Data** - if any data was missing or incorrect, the chatbot prompted the user to correct or provide the missing information.

## Data Mining Success Criteria

We ensured the success of our data mining process by:

- **Accurate data handling** - ensuring all data was correctly mapped to the premium and indemnity calculation logic.
- **Effective input validation** - ensuring that no erroneous data led to inaccurate results.
- **Error-free premium calculations** - all calculated premiums were based on valid inputs.

## Data Mining Goals

Our primary goal was to build a reliable and efficient system for transforming raw user inputs into actionable calculations that would generate accurate insurance quotes.

# Data Preparation

## Loading Data

Data was loaded directly from user input through the chatbot interface. Each piece of data was stored as a slot value, which ensured consistent and easy access during the calculation phase.

## Cleaning Data

We implemented automated data validation checks to ensure data quality:

1. **Numeric validation** - ensuring that numeric inputs, such as indemnity amounts were valid and within specified ranges.
2. **Completeness check** - ensuring that no required fields were left empty.
3. **User interaction** - if any input was invalid or missing, the chatbot prompted the user to correct or complete the required information.

## Analysis

Key Steps in Analysis:

1. **Staff Fee Calculation** - we calculated staff fees based on the number of partners, qualified/unqualified assistants, and other staff members.
2. **Annual Fee Calculation** - we applied a sliding scale to determine the annual fee based on the indemnity amount.
3. **Indemnity Limits Calculation** - we used multipliers to calculate the limit of indemnity based on the provided indemnity amount.
4. **Profession Fee** - we categorized business types and applied profession-specific multipliers to determine additional fees.
5. **Levies** - we factored in a levy fee to calculate the total premium.
6. **Premium Calculations** - we summed up staff fees, annual fees, professional fees and levies to determine the basic premium.

The analysis process ensured that all calculations followed a logical, error-free sequence, producing accurate and reliable results.

# Recommendations

Based on our analysis, we made the following recommendations:

1. **Improve Data Validation** - we recommended incorporating additional checks for edge cases in user input to ensure all possible scenarios were handled smoothly.
2. **Broaden Profession Coverage** - we suggested adding more business professions to the system to accommodate a wider range of users.
3. **Automate PDF Generation and Distribution** - we recommended implementing a fully automated system for generating and sending PDF quotes to clients.
4. **Explore Machine Learning for Risk Assessment** - we proposed introducing machine learning algorithms to assess risk dynamically and adjust premiums based on user data trends.