ACTL 4305/5305 Actuarial Data Analytic Applications

Assignment 2 (Group Assignment): Modelling and Evaluation

Due date: 4pm 20 November 2020 (Friday)

Learning outcomes

The assignment aims at developing the course learning outcomes in relation to modelling principles, predictive models, model selection and assessment, business and communication. It also assesses the program learning outcomes *Knowledge*, *Problem solving and critical thinking*, *Application of knowledge and skills*, *team work* as well as *Communication*.

Team Formation

In this project, you and your classmates will form a general insurance pricing team. You are expected to form your own teams (team size: 4-6) and submit your team formation form via the *'Assignment 2 Group Self-selection'* tool on moodle.

You can form a team with students from or outside of your tutorial. However, postgraduate students are expected to form a team with other postgraduate students only.

You may use the *[Assignment 2 Group Selection]* channel to assist with team formation (e.g. to recruit members for your team or to request to join a team).

You are strongly recommended to form your own teams by 30 October. Students who fail to do so by the team formation Deadline **30 October** will be randomly allocated to teams.

Tasks

As a team of pricing analysts for a general insurance company, you are provided with the dataset `A2-data.csv`, which includes the private automobile policy and claims information of 40,621 policies during the calendar years 2015 to 2018. The descriptions of variables in the dataset are provided in the next section below. One record (row) in the dataset can represent multiple claims from one insured.

Your team's task is to find out the best model for pure premium predictions (please refer to the section 'General Insurance Pricing: Pure Premium' below for the 'pure premium' definition). You are required to only consider models discussed in this course.

The data has already been cleaned and checked quality by the data exploration team. No more cleaning or quality check is needed. However, you are suggested to do some basic data exploration to understand and prepare the data before the formal modelling process.

To complete this task as a team, each team member is required to select one model for modelling and evaluation purposes. It is OK to have multiple team members working on the same model structure (but from different perspectives). And it is recommended to have a good spread of models explored by different members in the team. Each team should have good communications at the beginning of the project to determine the allocation of models and how to combine the modelling results together for an overall comparison later. For

example, it may be a good idea to split the data into training and testing set at the beginning. And the testing set should only be used for model comparison purposes in the end. Each team member could use the same training set for modelling and the testing set for model performance assessment. Different modelling results from different team members could then be compared based on the performance on the same testing set. A best model should be suggested in the end. Each team member is expected to have a similar level of involvement and contribution to the assignment.

To help you collaborate effectively in a team, you are required to draft a Team Contract and hold regular team meetings with meeting minutes. Team contract should include roles and responsibilities for each member and other team rules, such as meeting time, milestones and timeline. The minutes of team meetings must record chair and minute taker, and which members have been allocated which tasks. At each meeting each member should be given a task on which they must report progress at the following meeting. You are strongly recommended to read the "How to collaborate in a team effectively?" section on Moodle for more references and resources of teamwork. It introduces the importance of teamwork, which provides the motivation of having a group-based assessment task in this course. It also provides useful instructions on topics such as developing team skills, initiating a team's contract, establishing roles, working in virtual teams, etc.

Your teams tasks are listed below:

1. **Video Presentation**: Your team manager asked your team to prepare a *no more than 15 minutes* video presentation along with slides (each team member is expected to present) to report your project progress and outcome. It is suggested to include (but not limited to) the background, overall problem-solving strategy, the modelling process, model comparisons, model selection and conclusion. Please note your manager is not interested in technical details, but the problem solving process and results. For resources and recommendations on video production, please refer to the "Assignment 2" section on Moodle.
2. **Individual Modelling Report** Another pricing analyst, who has taken ACTL4305/5305 (very familar with the course content), will review your individual modelling process and results. Each team member should submit an individual modelling report using R Markdown including (but not limited to) a brief data exploration and preparation section, the model fitting and checking processes, model performance assessment (on the testing set), potential advantages and disadvantages of using this model. The report should include R coding details and outputs. There is no need to include other members' modelling outcomes in this individual report. The Individual modelling report should include *adequate but not excessive* information to help the review analyst understand your modelling process and results in a short period of time. There is no page limit of the technical report.
3. **Group-specific Discussion Forum**: Your team is required to post a team contract, minutes of each group meeting, individual posts and feedback to your team members' posts on the newly established A2 Discussion Forum for your group, which can be found in the "Assignment 2" section on Moodle. This is a formative assessment (similar to our weekly discussion forum) and is designed to help you collaborate in a team efficiently.
    - Please initiate a Team Contract as soon as possible (maybe after the first teams meeting) and post it on the A2 Discussion Forum for your team.
    - Each team must have at least three group meetings before the assignment due. The minutes of each meeting are required to be posted on the Discussion Forum on a timely basis.
    - Team members should also post frequently (at least once a week) on the discussion forum to report progress, ask questions/feedback and reply to other team members' posts with constructive suggestion and feedback before the assignment due date. Active participation and constructive feedback to each other's work are strongly recommended.

Variable Descriptions:

All the variables in the dataset can be categorised as continuous or categorical.

- index - the row ID.

1, 2, 4, 7, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 , 22 , 24 , 25, 26

- year - Calendar year, i.e., 2015, 2016, 2017, 2018.
- exposure - The time a car was exposed to risk during the calendar year.
- business.type - New business (nb) or renewal business (rb).
- driver.age - Age of driver.
- driver.gender - Gender of driver.
- marital.status - Marital status.
- yrs licensed - The number of years that the operator has been licensed.
- ncd.level - No claim discount level (higher level reflects a greater discount for not having any claims).
- region - The garaging location of the vehicle. The variable has been coded as a positive integer. We do not have any information about how these regions are spatially related.
- body.code - The body style of the vehicle.
- vehicle.age - The age of the vehicle.
- vehicle.value - The value of the vehicle.
- no.seats - The number of seats of the vehicle.
- cubic.cent - The size of engine in cubic centimeters.
- horse.power - The horse power.
- weight - The weight of the vehicle
- length - The length of the vehicle
- width - The width of the vehicle
- height - The height of the vehicle
- fuel.type - The type of fuel (gasoline, diesel, or liquefied petroleum gas)
- prior.claims - The number of prior claims.
- claim.count - Number of claims.
- claim.incurred - Sum of the individual claim payments (ultimate costs of all claims).

General Insurance Pricing: Pure Premium

If the number of policies in a collection, n, is large, then the average provides a good approximation of the expected loss

$$\mathrm{E}(X) \approx \frac{\sum_{i=1}^{n} X_i}{n} = \frac{\text{Loss}}{\text{Exposure}} = \text{Pure Premium}.$$

We define the 'pure premium' to be the sum of losses divided by the exposure; it is also known as a 'loss cost'.

We can multiply and divide by the number of claims, `claim count`, to get

$$\text{Pure Premium} = \frac{\text{claim count}}{\text{Exposure}} \times \frac{\text{Loss}}{\text{claim count}} = \text{frequency} \times \text{severity}.$$

So, when premiums are determined using the pure premium method, we can either take the average loss (loss cost) or use the frequency-severity approach.

Submission

Your Group's Final Solution Submission should include:

- Video Presentation. Format: mp4
- Slides for Video Presentation. Format: pdf.
- Individual Modelling Reports of All Members. Format: pdf
- Individual Modelling R Markdown Source File of All Members. Format: rmd. Only the PDF document will be marked, but the R Markdown file might be checked to clarify a response in the report if necessary.

You are required to upload all the documents/videos above in UNSW OneDrive and submit a link to your submission folder here. For more detailed instructions, please refer to the *Assignment 2 Submission* tool on Moodle. Please also refer to "UNSW Submitting a video via UNSW OneDrive" document on Moodle under the "Assignmetn 2" Section for instructions on how to share a folder link via UNSW OneDrive. Please also refer to other resources on Moodle for team work and video production.

Requirements

- *Due date:* Friday 20 November 2020, 4pm. There will be a 10 minutes grace period.
- *Late submission:* Please note that when an assessment item was submitted late, the School of Risk and Actuarial Studies will apply the following policy. A penalty of 25% of the mark the student would otherwise have obtained, for each full (or part) day of lateness (e.g., 0 day 1 minute = 25% penalty, 2 days 21 hours = 75% penalty).
- *Technical issues:* Students are reminded of the risk that technical issues may delay or even prevent their submission (such as internet connection and/or computer breakdowns). Students should then consider either submitting their assignment from the university computer rooms or **allow enough time** (at least 24 hours is recommended) between their submission and the due time. Please be reminded that the submission can be *slow* when it is close to the deadline if many students are submitting at a similar time.
- *Total marks:* 15 (15% towards the final grade)
- *Page limit:* There is no page limit.
- *Reference:* You do not need to reference any other material to complete this assignment but if you do, please ensure you properly reference your work.
- *Submitting documents:* Assignments must be submitted via the Assignment 2 submission box that is available on the course Moodle website.

Marking Criteria

Total marks: 15 marks (15% towards the final grade)

For group assessment items, each team member will share the same marks. For individual assessment items, each team member will be given individual marks.

1. Video Presentation: 35%
    - [Group assessment] Content: 15%
    - [Group assessment] Organisation: 10%
    - [Individual assessment] Presentation: 10%
2. [Individual assessment] Individual Modelling Report: 40%
    - basic data exploration and preparation (including the quality of communication): 10%
    - modelling process, such as feature selection, fitting, diagnostic checking and assessment (including the quality of communication): 30%
3. Group Discussion Forum 20%
    - [Group assessment] Each team needs to post a "Team Contract" (drafted after the first meeting) and minutes of each meeting (should have at least three meetings) on a timely basis (soon after the meeting): 10%.
    - [Individual assessment] Team members should post frequently (at least once a week) on the discussion forum to report progress, ask questions/feedback and reply to other team members' posts with constructive suggestion and feedback before the assignment due date: 15%.

Dealing with Group Issues and Conflict Resolution

Conflict is almost inevitable when you work with others. People have different viewpoints and, under the right set of circumstances, those differences may escalate to conflict. It is common that most groups experience issues at some time. What matters is how you handle that issue or conflict. This will determine whether it works to the team's advantage or contributes to its demise. Conflict is not necessarily a bad thing, though. Healthy and constructive conflict is a component of high-functioning teams. Conflict arises from differences between people; the same differences that often make diverse teams more effective than those made up of people with similar experience. When people with varying viewpoints, experiences, skills, and opinions are tasked with a project or challenge, the combined effort can far surpass what any group of similar individuals could achieve. Team members must be open to these differences and not let them rise into full-blown disputes.

Understanding and appreciating the various viewpoints involved in conflict are key factors in its resolution. These are key skills for all team members to develop. The important thing is to maintain a healthy balance of constructive difference of opinion and avoid negative conflict that is destructive and disruptive. One of the objectives of this project is to help you build communication and conflict resolution skills. The following resources are included to support you:

- Identifying Group Issues
- Dealing with Group Work Issues

Peer Review

There would be an option for your team to initiate a Peer Review Process for your group project. However, this should only be used as a last resort. You are strongly encouraged to negotiate a resolution yourselves within the team before initiating this formal review.

This is a completely optional process for conflict resolution. Mark adjustments may be applied according to the peer review outcome.

If your team would like to initiate a Peer Review Process, please contact Fei via Teams chat or Email.

Consultation and Advice

If your team needs consultation with the lecturer (Fei Huang). Please feel free to book a meeting with Fei via Teams chat or Email.