Predicting Court Decisions for Alimony: Avoiding Extra-legal Factors in Decision made by Judges and Not Understandable AI Models

Fabrice MUHLENBACH [1] Long NGUYEN PHUOC [2] Isabelle SAYN [3]

(1) Université de Lyon, UJM-Saint-Etienne, CNRS, Laboratoire Hubert Curien, UMR 5516, 18 rue du Professeur Benoît Lauras, F-42023 Saint Etienne, FRANCE
(2) Université Lumière Lyon 2, Maison des Sciences de l'Homme Lyon St-Étienne, 14 avenue Berthelot, F-69363 Lyon cedex 07, FRANCE
(3) CNRS, Université de Lyon, Centre Max Weber, UMR 5283, Maison des Sciences de l'Homme Lyon St-Étienne, 14 avenue Berthelot, F-69363 Lyon cedex 07, FRANCE

Is it possible to develop a method to predict the spouse alimony after a divorce?

In France, the law provides the transfer of income between spouses to tend to reduce inequality in living standards following a divorce by a list of non-exhaustive and non-prioritized criteria, making alimony one of the elements of the divorce decision the most difficult to anticipate.

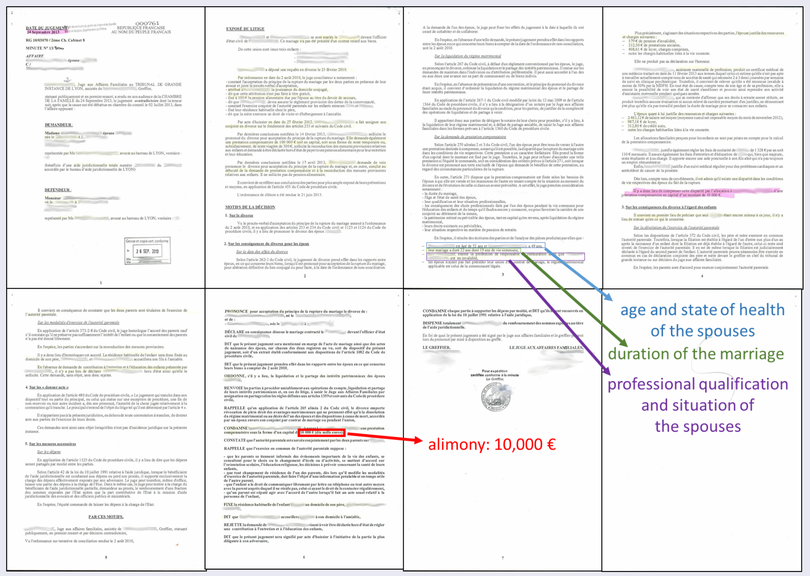
A model for predicting spouse alimony? What for?

Interest of various stakeholders:

- **Litigants** want to have an idea of what they can expect from the court decision (plan for the future, base for further negotiations...)
- **Lawyers** want to respond to the predictability concerns of their clients and to be able to establish a judicial strategy either by helping the divorcees to reach an agreement, or by making a legal claim.
- **Judges** are divided between protecting their appreciation and producing comparable decisions for cases of comparable clients, even more so for judges who assume managerial responsibilities within the jurisdiction.
- **Jurisdictional organization** considers predictability as the means to favor agreements and therefore to unclog the courts.
- **Researchers** want to produce knowledge about the determinants of court decisions, looking for decision-making mechanisms and knowing the determinants of court decisions, not only the legal one but also the others which interfere with the decision (bias or extra-legal factors).

Initial work: Manual analysis

Anonymized divorce court decision with an alimony amount and relevant variables:



⇒ Manual analysis and recording of 14 groups of variables including the 3 categories:

1. Request for alimony from the parties,
2. Judge's decision on alimony payments,
3. Reasons for the decision whether or not to award alimony.

These variables were selected first on the criteria of the Civil Code (article 270 and 271): the disparity between the respective living conditions after divorce, the needs of the spouse and the resources of the other (now and in the predictable future), the duration of the marriage, the age and state of health of the spouses; their professional qualification and situation, the consequences of the professional choices made by one of the spouses during the common life for the education of the children, and the time that it will still be necessary to devote to it or to favor the career of his/her spouse to the detriment of his/her own; the estimated or foreseeable patrimony of the spouses –both in capital and in income– after the liquidation of the matrimonial regime; their existing and foreseeable rights; their respective retirement pensions situation. These variables were selected secondly on possible non-legal determinants of the decisions, as a fault-based divorce, the temporary support payments, or the sex of the judge.

Design of the alimony prediction model

- Step 1: Predict the alimony eligibility and acceptance by the court,
- Step 2: Predict the alimony amount set by the court,
- Step 3: Adjust the alimony amount from Step 2 by the outcome from Step 1.

Learning phase: The model is trained with court decisions in a supervised way by using two submodels:

- a **classification model** in Step 1,
- and a **regression model** in Step 2.

Prediction of the adjusted alimony amount by considering both its acceptance probability and its amount:

$$\widehat{y}_{alimony} = \widehat{y}_C \times \widehat{y}_R$$

where $\widehat{y}_{alimony}$ is the adjusted alimony amount and $\widehat{y}_C$, $\widehat{y}_R$ are respectively the predicted variables of the Classification model and the Regression model.

Experiments: Dataset

- 5,453 total divorce decisions
- ⇒ 3,203 cases with an alimony
- ⇒ 2,678 cases with an alimony approved by the court
- → (1) 1,524 cases where the parties have agreed (alimony = amount proposed by the parties)
- → (2) 1,257 cases where the parties did not agree (prediction of the alimony amount)

Experiments: Classification model (Random Forest)

List of the most important features in classification whether to grant alimony or not using Gini importance:

Variables	Gini
Activity status of the wife	19.9
Activity status the husband	15.6
Salary of the husband	30.5
Retirement pensions of the husband	13.4
Salary of the wife	26.1
Other income of the wife	10.6
Nb of children from the couple	16.2
Nb of adult children of the couple	13.7
Common life during marriage	21.4
Temporary support payments	25.0
Temporary allocation of domicile	10.7
Capital paid at once requested	33.7
Type of capital in cash requested	16.2
Type of capital in cash offered	18.5
Seat of First Instance Court	107.1

Experiments: Regression model

List of the most important features in regression using forward stepwise:

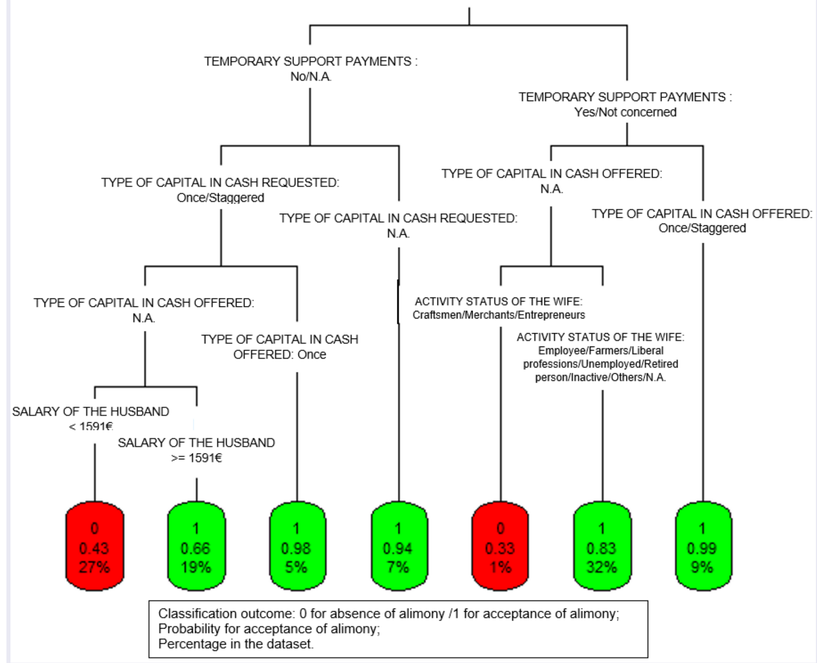
Variables	Estimate
Intercept	8403.15
Capital at once requested	0.33
Capital at once offered	0.79
Capital at once in a joint request	0.88
Capital over time offered	0.48
Capital over time in a joint request	0.80
Capital over time requested	-0.56
Months of capital over time requested	353.42
Pension offered	-88.72
Pension requested	40.80
Temporary pension offered	1.37
Salary of the wife	-7.86

Regression ⇒ $\widehat{y}_R$:

$$\begin{aligned}\widehat{y}_R = & \quad 0.33 \times \text{Capital at once requested} \\ & + \ 0.79 \times \text{Capital at once offered} \\ & + \ \dots \\ & + \ 1.37 \times \text{Temporary pension offered} \\ & - \ 7.86 \times \text{Monthly salary of the wife} \\ & + \ 8403.15\end{aligned}$$

General model

- Random Forest ⇒ $\widehat{y}_C$: [The figure below is an extract from a decision tree of the RF model]



- Regression ⇒ $\widehat{y}_R$
- General model to predict alimony amount: $\widehat{y}_{alimony} = \widehat{y}_C \times \widehat{y}_R$

Performance comparisons

Absolute errors in prediction (in thousands of euros):

Model	Mean	Median	σ	R^2
Ordinary least squares (OLS) Regression	21.46	10.64	35.93	0.66
Quantile Regression	19.73	9.04	40.93	0.62
RF×OLS Regression	16.46	3.95	35.59	0.70
RF×Quantile Regression	**15.95**	**3.43**	32.01	0.65

Discussion and conclusion

Objective: Test the possibility of designing a machine capable of following a reasoning allowing to reach a global result (a decision) or to know those of the criteria which determine the amounts retained by the magistrates, in the exercise of their discretion.

Usefulness: We know that judges use optional decision support tools to decide the amount, and we are not sure how these tools can influence judges' decisions. It is important to ensure the good quality of these tools.

Fairness: Discrimination between the determinants of decisions falling within the scope of the law and extra-legal determinants allows judges to assess the model and to keep the possibility of rejecting it. It is a means of respecting the rule of law and therefore democracy: the tools must respect the law and technical requirements must not prevail over the application of democratically voted laws.

Further work: We plan to develop automatic analysis systems for raw texts of court decisions in order to directly detect the values of interest in the text (with natural language processing and text mining techniques), avoiding the long and tedious phase manual document analysis.

Acknowledgments

This work is a part of *e-Juris* project, funded by the *Maison des Sciences de l'Homme de Lyon Saint-Etienne*.