

University of Groningen

## Feature selection and intelligent livestock management

Alsahaf, Ahmad

DOI:  
[10.33612/diss.145238079](https://doi.org/10.33612/diss.145238079)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2020

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*  
Alsahaf, A. (2020). *Feature selection and intelligent livestock management*. [Thesis fully internal (DIV), University of Groningen]. <https://doi.org/10.33612/diss.145238079>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



university of  
 groningen

# Feature Selection and Intelligent Livestock Management

**PhD thesis**

to obtain the degree of PhD at the  
University of Groningen  
on the authority of the  
Rector Magnificus Prof. C. Wijmenga  
and in accordance with  
the decision by the College of Deans.

This thesis will be defended in public on

Friday 4 December 2020 at 11.00 hours

by

**Ahmad Mohammed Jawad Ahmad Alsahaf**

born on 24 January 1989  
in Kuwait

**Supervisors**

Prof. N. Petkov  
Prof. R.F. Veerkamp

**Co-supervisor**

Dr. G. Azzopardi

**Assessment Committee**

Prof. A.C. Telea  
Prof. M. Biehl  
Prof. B. Kemp

---

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Machine learning in livestock science . . . . .	2
1.1.1	Estimated Breeding Values . . . . .	2
1.1.2	State of the art . . . . .	6
1.2	Feature Selection and model interpretability . . . . .	8
1.2.1	Sample re-weighting and feature importance . . . . .	10
1.3	Thesis Outline . . . . .	13
<b>I</b>	<b>Machine Learning And Computer Vision for Livestock</b>	<b>15</b>
<b>2</b>	<b>Phenotype Prediction: Slaughter Age in Pigs</b>	<b>17</b>
2.1	Introduction . . . . .	18
2.2	Materials And Methods . . . . .	18
2.2.1	Data . . . . .	19
2.2.2	Regression with RF . . . . .	19
2.2.3	Multiple Linear Regression . . . . .	20
2.2.4	Feature Importance . . . . .	21
2.2.5	Pedigree and Pedigree-Genetic Similarities . . . . .	21
2.2.6	Implications on Pen Assignment . . . . .	22
2.3	Results . . . . .	23
2.3.1	RF Regression Results . . . . .	23
2.3.2	Feature Importance . . . . .	23
2.3.3	Implication on Pen Assignment . . . . .	24
2.4	Discussion . . . . .	24
2.5	Conclusions . . . . .	30

**Appendices**

2.A	Supplementary results . . . . .	32
2.B	Assigning pigs to uniform target weight groups using machine learning	36
2.B.1	Abstract . . . . .	36
2.B.2	Introduction . . . . .	36
2.B.3	Materials and methods . . . . .	37
2.B.4	Results . . . . .	39
2.B.5	Discussion and conclusion . . . . .	40
<b>3</b>	<b>Muscularity Estimation of Pigs with Computer Vision</b>	<b>43</b>
3.1	Introduction . . . . .	44
3.1.1	Background . . . . .	44
3.1.2	RGB-D computer vision . . . . .	44
3.2	Data, Methods, and Experimental settings . . . . .	45
3.2.1	Data description . . . . .	45
3.2.2	Image pre-processing and selection . . . . .	46
3.2.3	Feature extraction . . . . .	49
3.2.4	Classification . . . . .	49
3.3	Results . . . . .	50
3.4	Discussion . . . . .	51
3.5	Conclusion . . . . .	55

**II Feature Selection 57**

<b>4</b>	<b>A Framework for Feature Selection Through Boosting</b>	<b>59</b>
4.1	Introduction . . . . .	60
4.2	Proposal and related work . . . . .	62
4.2.1	Boosting and feature selection . . . . .	63
4.2.2	Ensemble tree models and feature selection . . . . .	66
4.3	Methodology . . . . .	67
4.4	Experimental settings and evaluation . . . . .	67
4.4.1	Compared methods . . . . .	69
4.5	Results and discussion . . . . .	70
4.6	Conclusion . . . . .	73
4.A	Iterative Input Selection . . . . .	75
4.B	The effect of model evaluations on selection . . . . .	75
4.C	Weighting strategy and reset . . . . .	75
4.D	Validation with other classifiers . . . . .	76

*Contents*

*Contents*

---

<b>5 Outlook and conclusions</b>	<b>83</b>
<b>Publications</b>	<b>101</b>
<b>Summary</b>	<b>103</b>
<b>Samenvatting</b>	<b>105</b>



---

## Acknowledgments

Carl Sagan, astronomer and renowned science popularizer, once comically claimed that If one wished to make an apple pie from scratch, they must first invent the universe. Though his statement would not have led to a concise apple pie recipe, it held in its absurdity a deep truth. Even the most atomized and individualistic projects are predicated on countless foundations, many of which are invisible to the project's authors. Like Sagan's pie, the completion of this thesis is a project for which I cannot take full credit, nor claim to have "made from scratch". For this reason, I must be thorough in my declaration of gratitude. Not so thorough as to causally trace the atoms of this book back to the big bang, but enough to give some people the credit they deserve.

In the few days when I was inspired to go to the office at an early hour - by my standards anyhow - the 5<sup>th</sup> floor of the Bernoulliborg was quietly empty. Besides the occasional encounter with one of the restless researchers, whose early starts were evidently more habitual than my own, the only other faces I saw in those morning hours were those of the cleaning workers. When they saw me, they always greeted me with gentle smiles, and I happily requited.

I like to think that those encounters were snapshots from a parallel world, momentarily overlaid over ours; a world where the toils of manual labour, and the joys of scientific research, were spread more equitably among us. And those smiles were in solemn recognition of the collective pursuits we were all part of, those of happiness and the betterment of society. In those particular days, it was my turn to research, and theirs to keep house.

But we live in this world of divided labour, where one's vocation and their personhood are surgically fused, like two vertebrae that had lost their capacity to freely move. So until a better day comes, I have only my gratitude to offer them. Likewise, I thank the secretaries, the canteen workers, the persons who re-filled and



serviced the coffee machines, the grocery store workers, the baristas at the coffeehouses where I often worked, the people who paved the roads on which I cycled to work, and the welders who bound the bicycle's steel tubes together. The list could go on ad nauseam, but I cut it short, lest its length detracts from the sincerity of my thanks. You are encouraged, if you will, to extend the list in your own mind.

Due to the nature of this PhD project, I must give special thanks to the pig farmers and handlers who reared the 30,000+ animals that ended up as data points in my thesis. Those farmers, a few of whom I've had the bitter-sweet pleasure of seeing in action, were more diligent in their work than I could have been in mine. While I sat at my desk, breathing the cool, mundane air that circulates through an office building, they worked at those dimly lit porcine gulags; their noses getting slowly adapted to the odors of death, despair, and excrement. If not for olfactory fatigue, I would have thought that theirs was a fate no better than the pigs'.

The limits of my work difficulties were a few stubborn software bugs, and some similarly stubborn peer reviewers. The farmers, meanwhile, bore on a daily basis the duties of feeding the pigs, cleaning their manure, tagging their ears, docking their tails, castrating the males, forcibly inseminating the females, and eventually, sending them off to slaughter. I was there to observe them for just one day, but their competence does not need my testimony. A visit to the perpetually stocked meat aisle at your nearest grocery store will suffice.

Grim as it may be, there is an elegance to how slaughtering one of earth's most intelligent species has been industrialized. A single well-oiled machine that turns the animated snorting and squealing bodies into carcasses.

Though it wasn't all in vain. Those carcasses ended up fragmented on dining tables around the world, hopefully satisfying many pallets, and bringing joy to those least aware from whence their meals came. For me, they did a tad bit more. They may well have laid the basis for my forthcoming livelihood. It's as if they were 30,000 fleshy platforms in a Nintendo game, and I was the mustachioed plumber jumping across them to get to the other side. Not to Princess Peach's castle, but to a middle class life, and a better chance at self-actualization. Thereby avoiding a pit of uncertainty that I could have fallen into, had I not been blessed with just the right circumstances.

Maybe that's enough solace for the pigs.

If a copy of this book survives to a time when humans no longer enslave and kill the innocent for profit, I hope whoever lifts it from the rubble and reads these words will understand and forgive us, for we knew not what we were doing. Those farmers, the breeders, the butchers, and myself; We were just doing our jobs.

To my colleagues at the University of Groningen, and all researchers who would consider themselves my peers. From the most uncompromising arbiters of scien-

tific truth to the most cynical careerists, and everyone else in between. You are all my comrades, and I am honored to have been considered of your ilk, even if just temporarily. If I am fortunate enough to continue being an academic, I hope I will live up to higher standards of scientific integrity and intellectual responsibility than I have thus far.

To my advisors and mentors. Thank you for your support and your patience.

To my friends and family. Thank you for your love and sacrifice. I hope I did not let you down.



---

## List of Figures

1.1	Synthetic dataset showing the effect of sample re-weighting . . . . .	11
2.1	PCA explained variance . . . . .	24
2.2	Feature importance for slaughter age prediction . . . . .	26
2.3	Relative importance of data types in slaughter age prediction . . . . .	27
2.4	Fidelity of random projection . . . . .	33
2.5	Decision-Tree visualization (All data) . . . . .	35
2.6	Decision-Tree visualization (EBVs) . . . . .	35
2.7	Multiple instances of random forest feature ranking. . . . .	40
3.1	Color and depth images of the pigs . . . . .	46
3.2	Images of pigs of different muscularity . . . . .	47
3.3	The distribution of pig liveweight . . . . .	47
3.4	Depth image processing procedure . . . . .	48
3.5	Images of pigs with different postures . . . . .	49
3.6	Muscularity classification confusion matrix . . . . .	51
3.7	Input feature importance for muscularity classification . . . . .	52
4.1	Accuracies of selected subsets validated with a KNN classifier . . . . .	72
4.2	The effect of model evaluations on subset accuracy . . . . .	76
4.3	The effect of the reset strategy on the selected subsets . . . . .	77
4.4	Accuracies of selected subsets validated with an XGBoost classifier . . . . .	78
4.5	Accuracies of selected subsets validated with a Naive Bayes classifier . . . . .	80
4.6	Comparison of sample weighting strategies . . . . .	81
5.1	The trend of publications on machine learning and livestock . . . . .	85



---

## List of Tables

1.1	An example of milk yield records . . . . .	4
1.2	Feature importance scores with redundancy . . . . .	11
1.3	Feature importance scores after sample re-weighting . . . . .	13
2.1	Regression performance of principal components . . . . .	25
2.2	Regression performance results for slaughter age prediction . . . . .	25
2.3	Full data description . . . . .	32
2.4	Supplementary regression results . . . . .	34
2.5	Full data description: pen assignment . . . . .	38
2.6	Pen assignment classification error . . . . .	39
3.1	Input variables for muscularity classification . . . . .	50
4.1	Descriptions of datasets used in the comparison. . . . .	71
4.2	Computational time of the compared methods . . . . .	73



*"Whatever does not spring from a man's free choice, or is only the result of instruction and guidance, does not enter into his very being, but remains alien to his true nature; he does not perform it with truly human energies, but merely with mechanical exactness..  
..we may admire what he does, but we despise what he is."*

Wilhelm von Humboldt