Technical due diligence report for early stage investors

Company: X
From: Thomas Wood, Director: Fast Data Science Ltd

To: X

Date: X

This template is designed for technical due diligence on companies using machine learning. It does not include the financial questions that form the main part of a due diligence investigation, nor does it address the marketability of a product.

Executive summary

Architecture and code

Is the software documented?	
Code maintainability	
Unit tests	
Version control and ticketing	
Look in Git: who has committed and how frequently?	
Is it possible to understand the code with no introduction?	
Legacy or obsolescence?	
Remarks	

Machine learning model

Is the model accurate enough for its intended purpose? Does it work?	
Is the machine learning model under version control?	
How will the model be retrained?	
What KPIs will be used to score the model?	
Will there be a closed loop of data coming in for retraining after deployment?	
If so, how will future retrained models be quality controlled?	
Is AI bias likely to be an issue?	
Accuracy, precision/recall, ROC, AUC etc?	
Does the model output a degree of confidence in its predictions?	
Does the model require fallback to human intervention? How is this fallback decided?	
Remarks	

Data processing

Is the data processed in the cloud or on premise?	

Remarks	

Data

Has enough data already been collected?	
Cold start problem?	
Will GDPR/data protection be an issue?	
Is data collected so far specific to one customer or location and unlikely to generalise to new situations?	
Is there a bottleneck in data gathering if humans are needed to label every instance?	
Remarks	

Scalability

Will the services need to be separated across servers?	
Will sessions be stored?	
How will costs scale as the operation scales? (Pay special attention to cloud computing costs)	
Remarks	

Intellectual property

Is the technology patented?	
Is the technology easy to copy?	
Is it hard to obtain data to train this kind of model?	
Are there any limitations in terms of licenses of software or patents that are being used?	
Does the technology depend on third-party or open-source code, model architecture, dataset, or model weights?	
Was transfer learning used to develop the model, or was it trained from scratch?	
Any vendor lock-in?	
Remarks	

People

Who are the most important members of the team?	
Are their credentials and experience valid as claimed?	
What are their engagements? Full time, part time?	
What are their responsibilities?	
Who is engaged with this company as their primary activity?	
Is anybody crucial to the organisation? Any bus factor of 1?	

Any employee with e.g. exceptional talent for writing code?	
Any skills gaps?	
Any redundancies? Do team's skills complement each other or do they overlap (e.g. 100% PhD mathematicians or 100% developer background)?	
Remarks	

Product support

How will tech issues be resolved in the future?	
Remarks	

Regulation

Does it need regulation of any kind?	
Are we dealing with personal data?	
Do we need to consider environmental regulations? (if applicable)	
Is it licensed as a medical device? (if applicable)	
What needs to be done to be able to sell it as a **non-medical** device? (if applicable)	
What additional approval would be needed to be able to sell it as a **medical** device? (if applicable)	
Remarks	

Future roadmap of the company

What are the plans for the next few years?	
Remarks	

Risks

What technical risks are possible?	
Remarks	