



<http://www.aimspress.com/journal/dsfe>

Research article

Classifying high-potential startups for strategic partnerships using machine learning – The case of german digital startups

Dung Hai Dinh^{1,*}, Van Thanh Tran¹ and Nicole Ondrusch²

¹ Vietnamese-German University, Vietnam

² Heilbronn University, Germany

* **Correspondence:** Email: dung.dh@vgu.edu.vn.

Appendix

Appendix A1. Demographic characteristics of variables.

Variables	Number	Percent (of total startups)
Industry		
AdTech	16	2
Cybersecurity	58	8
Digital Health	126	17
Digital Chemistry	37	5
E-Commerce	71	9
Education	58	8
FinTech	112	15
InsurTech	86	11
LegalTech	24	3
Logistics	157	21
MediaTech	66	9
Mobility	130	17
SaaS	315	42
Smart Infrastructure	141	19
Smart Systems	177	23
Cross-industry Technology	121	16
Artificial intelligence	366	48

Blockchain	69	9
Robotics	46	6
Virtual Reality	49	6
Hardware	114	15
Software Development	486	64
Data Analytics	352	47
Internet of Things	197	26
Looking for		
Partner	617	82
Financing	451	60
Talents	411	54
Mentoring	331	44
Funding phase		
Pre-Seed	74	10
Seed	225	30
Early Stage	182	24
Growth Stage	163	22
Later Stage	7	1
Hub Affiliation		
Hub Berlin	77	10
Hub Cologne	49	6
Hub Dresden/Leipzig	69	9
Hub Dortmund	16	2
Hub Frankfurt/Darmstadt	142	19
Hub Hamburg	68	9
Hub Karlsruhe	54	7
Hub Mannheim/Ludwigshafen	85	11
Hub Munich	71	9
Hub Potsdam	46	6
Hub Nuremberg/Erlangen	46	6
Hub Stuttgart	25	3
Not part of the network yet	5	1
Size		
1-10	499	66
10-50	231	31
50-100	15	2
>100	10	1
Market		
Germany	108	14
DACH	114	15
Europe	139	18
International	320	42
Business Type		
b2b	699	93
b2c	231	31
Distance (category)		
Very Close (0-20 km)	84	11
Close (20-50 km)	57	8
Commutable (50-200 km)	189	25
Far (>200 km)	425	56

**Percentages may not add up to 100 as some variables are multi-label or contain missing values.*